

GenCore version 5.1.4 PJ5\_4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 8, 2003, 05:00:53 ; Search time 153 Seconds

(without alignments)  
7189.336 Million cell updates/sec

Title: US-09-600-714-41

Perfect score: 1254  
Sequence: 1 atgagctcctaagtaaccgcg.....attgagctgtgattttaa 1254

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 593429 seqs, 438583890 residues

Total number of hits satisfying chosen parameters: 1186858

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications NA:  
1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq:\*  
2: /cgn2\_6/ptodata/2/pubpna/PCF\_NEW\_PUB.seq:\*  
3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq:\*  
4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq:\*  
5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq:\*  
6: /cgn2\_6/ptodata/2/pubpna/PCFUS\_PUBCOMB.seq:\*  
7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq:\*  
8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq:\*  
9: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq:\*  
10: /cgn2\_6/ptodata/2/pubpna/US09\_PUBCOMB.seq:\*  
11: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq:\*  
12: /cgn2\_6/ptodata/2/pubpna/US10\_PUBCOMB.seq:\*  
13: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq:\*  
14: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	165.4	13.2	1440	10	US-09-949-145-5
2	165.4	13.2	1952	10	US-09-949-145-1
3	150.2	12.0	1497	10	US-09-949-145-6
4	150.2	12.0	2097	10	US-09-949-145-2
5	139.8	11.1	1486	10	US-09-864-761-1547
6	111.8	8.9	123	10	US-09-864-761-18306
7	109.6	8.7	1650	9	US-10-012-542-75
8	56.6	4.5	1446	10	US-09-864-761-20691
9	48.6	3.9	1973	10	US-09-864-761-3411
10	48.6	3.8	401	10	US-09-864-761-13936
11	47.4	3.8	570	10	US-09-864-761-21778
12	47.4	3.8	1976	10	US-09-864-761-5052
13	46.6	3.7	1403	10	US-09-864-761-19241
14	46.6	3.7	1403	10	US-09-864-761-19241
15	43.8	3.5	1635	10	US-09-864-761-20241
16	43.8	3.5	155074	9	US-10-026-188-6
17	43.6	3.5	4339	10	US-09-864-761-20174
18	42.4	3.4	574	10	US-09-864-761-228
19	42.4	3.4	669	10	US-09-864-761-17051

20	42	3.3	283	10	US-09-864-761-17994	Sequence 17994, A
21	41.4	3.3	370	10	US-09-728-446-223	Sequence 223, App
22	41	3.3	331	10	US-09-864-761-17053	Sequence 17053, A
23	41	3.3	465	10	US-09-864-761-230	Sequence 230, App
24	41	3.3	493	10	US-09-864-761-5863	Sequence 5863, App
25	41	3.3	6909	9	US-09-975-719-111	Sequence 111, App
26	40.6	3.2	201	10	US-09-864-761-22210	Sequence 22210, A
27	40.6	3.2	477	10	US-09-864-761-5436	Sequence 5436, App
28	40.4	3.2	830	10	US-09-864-761-19531	Sequence 19531, A
29	40.2	3.2	459	10	US-09-864-761-1233	Sequence 1233, App
30	39.6	3.2	1959	10	US-09-864-761-4012	Sequence 4012, App
31	39.4	3.1	276	10	US-09-864-761-25120	Sequence 25120, A
32	39.4	3.1	537	10	US-09-864-761-8390	Sequence 8390, App
33	39.2	3.1	583	10	US-09-864-761-20772	Sequence 20772, A
34	38.8	3.1	249	10	US-09-864-761-16846	Sequence 16846, A
35	38.8	3.1	476	10	US-09-864-761-6	Sequence 6, App1
36	38.8	3.1	479	10	US-09-864-761-768	Sequence 768, App
37	38.4	3.1	456	10	US-09-864-761-4249	Sequence 4249, App
38	38.2	3.0	2712	10	US-09-919-172-40	Sequence 1218, App
39	38	3.0	1200	9	US-09-738-626-1218	Sequence 142, App
40	37.8	3.0	500	9	US-09-854-133-142	Sequence 19900, A
41	37.8	3.0	500	10	US-09-738-973-142	Sequence 142, App
42	37.8	3.0	532	10	US-09-864-761-19900	Sequence 19900, A
43	37.8	3.0	58985	9	US-09-901-152-3	Sequence 3, App1
44	37.6	3.0	968	10	US-09-864-761-9629	Sequence 9629, App
45	37.4	3.0	350	10	US-09-864-761-23424	Sequence 23424, A

ALIGNMENTS

RESULT 1									
US-09-949-145-5									
Sequence 5, Application US/09949145									
Patent No. US20020055622A1									
GENERAL INFORMATION:									
APPLICANT: New York Blood Center									
TITLE OF INVENTION: Mammalian No. US20020055622A1-erythroid Rh Type C Genes and Glyc									
FILE REFERENCE: Docket 454-31									
CURRENT APPLICATION NUMBER: US/09/949, 145									
CURRENT FILING DATE: 2001-09-07									
PRIOR APPLICATION NUMBER: US 60/230660									
PRIOR FILING DATE: 2000-09-07									
NUMBER OF SEQ ID NOS: 77									
SOFTWARE: PatentIn version 3.1									
SEQ ID NO 5									
LENGTH: 1440									
TYPE: DNA									
ORGANISM: Homo sapiens									
US-09-949-145-5									
Query Match									
Best Local Similarity 49.1%; Pred. No. 6.2e-42;									
Matches 439; Conservative 0; Mismatches 456; Indels 0; Gaps 0;									
QY	141	CTATCAAGTTGGCCAAAGTCTGACGCTGAGGCGGCAATGGCTTGAGGCTTCTCACCTC	200						
DB	177	CTACCAAGCTTCCAGACGTCGACGTCGATGCTTCTTGCGGCTTCGCTTCTCAATAC	236						
QY	201	GAGTTTCCGAGACACAGCTGAGCAGTGTGACCTTCAACCTTTCATGCTGAGGCTTGG	260						
DB	237	TTTCTGAGGCGTACGAGCTTCAAGGCTTCAAGGCTTCAAGCTTCTGTTGAGGCTTGG	296						
QY	261	TGTGCAATGGGCAATCTGCTGAGACGCTTCTTCAAGCAGTCTTCTTGTGGAAGTGT	320						
DB	297	CATCAAGTGGGCGTGTCTCATGACGAGGCTGTTCACCTTCTTAAGACCGCTCATGT	356						
QY	321	CATCAAGTGTTCAGTATTCGCTGAGCAGCATATGATGCTTGTGCTGATCTCAGT	380						
DB	357	CGTGGGCTGAGAACCTCATCAAGCTGACTTCTGCTGAGCTTCTGTCTGCGGCTT	416						
QY	381	GGATGCTGTCTTGGGGAAGTCACTTGGCGCAGTGTGTGTGATGTGTGTGAGGT	440						

Db	417	TGGGGCAGTTCTGGGTAAAGTCACGCCCATTCAGCTGCTCATATAGACTTTCTTCCAACT	476
Qy	441	GACAGCTTTAGGCCAAGCTGAGGATGTCATCATGATATATCTTCAACAGACTACCAAT	500
Db	477	GACCCCTCTCCCTGCTGATGATGATTCATTTCTTAACTCTCTAAAGGTGAGATCCAG	536
Qy	501	GACATGATGACATCTACGTTTCGAGCCATTTTGGCTGTCTGTGGCCGTGGCCCT	560
Db	537	AGGCTCCATGACATTCACACATTTGGGGCTTACTTTGGGCTCACAGTACCCGGATCT	596
Qy	561	GCCAAAGCCTTACCCGAGGAAACGAGATTAAGATCAGACAGCAACGATCCCACTT	620
Db	597	CTACCGACGCAACTTAGAGCGAGACAGACAGACAGAAATTTCTGTGATCCAGTCGACCT	656
Qy	621	GTCCTGCATGCTGGGGCCCTCTTCTTGTGTGATGTTTGGCCAAATTTCACCTCTCTCT	680
Db	657	CTTTGCATGATTTGGACCCCTCTTCTGTGTGATGATCTGGCCCAAGCTTCAACTCAACCAT	716
Qy	681	GCTGAGAACTCATTCGAAAGAAAGATGCGGTTCACACCTCATATGCTGTACAGT	740
Db	717	ATCTCAATGAGGAGACAGCCAGCACGAGCCGCGCATTAACCTACTGCTCTTGGCAAC	776
Qy	741	CAGCTGTGTGACAGCCCATCTCAGGGTCATCTTGGCTCACCCCCAAGGAGATCAGCAA	800
Db	777	CTCGGTGCTTACCTCGGTGGCAATATCAAGTCCCTGACACAGAAAGGGCAAGCTGAGCAT	836
Qy	801	GACTTATGTGACAGTGGGGTGTGTGGCAGAGAGGTGGCTGTGGGATACCTCGTGTCACT	860
Db	837	GGTGCAATCCAGAAATGCAAGCTCGAGAGAGAGGGGTGGCCGTGGGATACCGCTGTCAGAT	896
Qy	861	GATCCCTCTCCGTGGCTGTCATGATGATGTGGGTCTTGTGGCTGGGTGATCTCCCTCG	920
Db	897	GATGTCATGCTTAAACGTGCTTCATCATTCGCGTTGTCTGGGGCATATATCTCCACCT	956
Qy	921	GGGAGCCAAATGACTGCGGGGGTGTGTAACGAGTCTGGGATTTCCCAAGCTTCAT	980
Db	957	GGGTTTGTATACCTGACCCCATTCCTGAGTCCCGGCTCAACATCCAGACACATGTGG	1011
Qy	981	CATGGGCTACAACTTCAGCTTGCGGGGTCTTGGAGAGATCATCTACATTTTG	1035
Db	1017	CATTACATCTGCATGGCATTTCTGCGATATATAGGCGCATGTTGGGTCTGTG	1071

```

RESULT 2
US-09-949-145-1
/ Sequence 1, Application US/09949145
/ Patent No. US20020055622A1
/ GENERAL INFORMATION:
/ APPLICANT: New York Blood Center
/ TITLE OF INVENTION: Mammalian No. US20020055622A1-erythroid Rh Type C Genes and Glyco
/ FILE REFERENCE: Docket 454-31
/ CURRENT APPLICATION NUMBER: US/09/949,145
/ CURRENT FILING DATE: 2001-09-07
/ PRIOR APPLICATION NUMBER: US 60/230660
/ PRIOR FILING DATE: 2000-09-07
/ NUMBER OF SEQ ID NOS: 77
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 1
/ LENGTH: 1952
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ PUBLICATION INFORMATION:
/ DATABASE ACCESSION NUMBER: AF193809
/ DATABASE ENTRY DATE: 1999-12-22
/ RELEVANT RESIDUES: (1)..(1952)
/ US-09-949-145-1

```

	Query Match	13.2%	Score 155.4	DB 10	Length 1952
	Best Local Similarity	49.1%	Pred. No. 7.3e-42		
	Matches 439	Conservative 0	Mismatches 456	Indels 0	Gaps 0
QY	141 CTATCAAGTGGCCCAAGATCTGACCGTATGAGCGGCATTGGCTTGGGCTTCTCACCCTC	200			

Db	201	CTTCCCAAGCTTCCAGGACGTCAGCGTAGATGCTTCTGTGGGGCTTGGGCTTCTCATTAC	260
OY	201	GAGTTTCCGGAGACACAGCTGAGAGCACTGTGGCTTTCAACTCTTACCTGTGGCGCTTGG	260
Db	261	TTTCTCGAAGGCTACAGGCTTACAGCGCGGTGGGCTTCAACTCTGTGTGGACACTTGG	320
OY	261	TGTGCAAGGGGCANATCCGTGCTGAGCGGCTTCTTAGCGAGTCCCTTCTGGGAAGGTGT	320
Db	321	CATCAGAGGGCGGCTGCTCATGCAAGGCTGGTTCCACTTCTTACAAACCGCTACATCGT	380
OY	321	CATCACTGTTCAGTATTTGGGCTGGCCACCATGATGCTTTGTGGTGTGATCTTCAGT	380
Db	381	CGTGGGCGCTGGAGAACTCATCAACGCTGACTTCTGCGTGGCCCTGTCTGCGGTGGCTT	440
OY	381	GGATGCTGTCTTTGGGGAAAGGTGCATCTTGGGCGAGTTGGGTGTATGTGTGCTGGAGGT	440
Db	441	TGGGCACTTCTGGGTAAAGTCACAGCCCAATTCAGCTCTCATACATGACTTTTCCAAGT	500
OY	441	GACAGCTTTAGGCAACCTGAGAGATGTCATCAGTAAATATCTTCAACACAGACTACCAT	500
Db	501	GACCTCTTCTGCTGATGATAGTTCATTCTTCTTAACTCTGTAAAGGTGAAGATGAGG	560
OY	501	GAACTATGTCACATCTPACTGTTCGACGCTTATTTGGGCTGTCTGTGGCTGTGGTCTT	560
Db	561	AGGCTCCATGACATTCACACATTTTGGCGCTTACTTTTGGGCTCACAGTGCACCGGATCT	620
OY	561	GCCAAAGCTCTTACCCGAGGGGAAAGAGATTAAGATCAGACAGCAACGATCCCACTT	620
Db	621	CTACCGACGCAACTTAGACAGACAGCAAGSAGACAGAAATTTGTGTACCAAGTGGACT	680
OY	621	GTTGCCCATGCTGGGCGGCTCTTTCTTGTGTAGTTCTGSCCAATTTCACTCTGTCTT	680
Db	681	CTTTGCCATGATGTGGACCCCTCTTCTGTGGATGTACTGGCCCGACGCTTCAACTAGCCAT	740
OY	681	GCTGAGAGTCCATTCGAAAGSAGAAATGTCGTGTAAACCTACTATGCTGTAGAGT	740
Db	741	ATTCCTACCATGGGAGACAGCCAGCACCGAGCGCCATTAACACCTACTGTCTTGGAGC	800
OY	741	CAGCTGTGTGACAGCCCATCTTCAGGGTCACTCTTGGCTCACCCCCAAGSAGATTCAGCA	800
Db	801	CTGGGTGCTTACCTCGGTGGCAATATTCAGAGCCCTGCACAAAGSAGGCAAGCTGAGCAT	860
OY	801	GACTTATGTGACACAGTGGGCTGTGTGGCAGAGGCGTGGCTGTGGGTACTCTGTGTCACT	860
Db	861	GGTGACATTCAGAAATGCAAGCTCGAGAGAGGGGTGGCGGTGGTACCCCTGCTGAT	920
OY	861	GATCCCTTCTCGAGGCTGTCATGATGTGTGGGTCTTGTGGCTGGGTGATATTCCTCCGCG	920
Db	921	GATGTCAATGCTTAAAGGTGCTTCATCATCGGCTTGTGTCTGGGCAATCATCTCACCTT	980
OY	921	GGAGACCAAGTACCTGCGGGGGTGTGTAACCGAGTCTGGGATTTCCCAACAGCTCAT	980
Db	981	GGGTTTGTATACCTGACCCCATCTCTGAGATCCCGGCTGCACATTCAGAGACATGTGG	1040
OY	981	CATGGGCTTACCAATTCAGCTTGTGTGGGTCTGCTTGGAGACATCATCTAACATTGG	1035
Db	1041	CATTAACTATCGATGCGCATTTCTGTGCATCATAGGGGCAATGCTGGTCTGTGG	1095

```

RESULT 3
US-09-949-145-6
; Sequence 6, Application US/09949145
; Patent No. US20020055622A1
; GENERAL INFORMATION:
; APPLICANT: New York Blood Center
; TITLE OF INVENTION: Mammalian No. US20020055622A1-erythroid Rh Type C Genes and Glyc
; FILE REFERENCE: Docket 454-31
; CURRENT APPLICATION NUMBER: US/09/949,145
; CURRENT FILING DATE: 2001-09-07
; PRIOR APPLICATION NUMBER: US 60/230660
; PRIOR FILING DATE: 2000-09-07
; NUMBER OF SEQ ID NOS: 77
; SOFTWARE: PatentIn version 3.1

```

```

; SEQ ID NO 6
; LENGTH: 1497
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-949-145-6

```

Query Match	12.0%	Score 150.2	DB 10	Length 1497
Best Local Similarity	47.5%	Pred. No. 4.2e-37		
Matches 446	Conservative	0	Mismatches 493	Indels 0
			Gaps	0

[illegible]

```

RESULT 4
US-09-949-145-2
; Sequence 2, Application US/09949145
; Patent No. US20020055622A1
; GENERAL INFORMATION:
; APPLICANT: New York Blood Center
; TITLE OF INVENTION: Mammalian No. US20020055622A1-erythroid Rh Type C Genes and Glyc
; FILE REFERENCE: Docket 454-31
; CURRENT APPLICATION NUMBER: US/09/949,145
; CURRENT FILING DATE: 2001-09-07
; PRIOR APPLICATION NUMBER: US 60/230660
; PRIOR FILING DATE: 2000-09-07
; NUMBER OF SEQ ID NOS: 77
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 2
; LENGTH: 2097
; TYPE: DNA
; ORGANISM: Mus musculus
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: AF193810
; DATABASE ENTRY DATE: 1999-12-22
; RELEVANT RESIDUES: (1)..(2097)
; US-09-949-145-2

```

Query Match	12.0%	Score 150.2%	DB 10%	Length 2097%
Best Local Similarity	47.5%	Pred. No. 5e-37		
Matches	446	Conservative	0	Mismatches 493
				Indels
				Gaps
				0
Qy	97	CACATATACGCTTCTTATAGAGATCAAAAAGGGCTCGTGCACTCTTATCAAGTTGGCCAA	156	
Dy	258	CGCAAGAACATCTCCACAGCGATTTGAACAGATTTTACTATCGCTACCCAGGCTTTCAG	317	
Qy	157	GATCTGACCGGATGAGGGCGGCATTTGGGCTTTCCTCACTCGAAGTTTCGGAGACAC	216	
Dy	318	GATGTACACGCATAGTCTTCTGTGGCTTGGGCTTCTCTCATGACCTTCTCGACGCTAC	377	
Qy	217	AGCTGAGACAGTGTGGCTTCAACCTTTCATAGCTGGCGCTTGTGTGACAGTGGCAATC	276	
Dy	378	GGCTTCAGCGCTGAGGCTTCACTTCTCGCTGACAGCTTTCGGCATTCACATGGGCACTG	437	
Qy	277	CTGTGTGACCGGCTTCTGAGACAGTTCCCTTCGGGAAAGTGTGTATCACACTGTTCAGT	336	
Dy	438	CTCATGACGGATGATTTCCATTTACTTTGAAAGAACCCACATTTGCTCGAGCTGAGAAC	497	
Qy	337	ATTGGGCTGGCCACATGAGTGTCTTGTTCGGTGTGATCTTCAGTGAATGCTGTCTTGGGG	396	
Dy	498	ATCATCAAGTGACTTCTGTGTGGCATCTTCTGTGTGGCTTTCGGGGCAGTTCTAGGC	557	
Qy	397	AAGGTCACTTTGGCGCAGTTGGTGTATGTGTCTGTGTGAGGTGACACTTTTAGGCAC	456	
Dy	558	AAGGTCACTTTGGCGCAGTTGGTGTATGTGTCTGTGTGAGGTGACACTTTTAGGCAC	517	
Qy	457	CTGAGATGATGATCAGTAATATCTTCAACACACACTAACACATGAACATGATGACATC	516	
Dy	618	AATAGTTCATCTCTCTGAACTCATAGAGCAAAAGTAGAGGGGCTCTATGACATTC	677	
Qy	517	TACGTGTTCGACGCTTATTTTGGGCTGTCTGTGGCTGTGGCTGTGCCAAGCCTTACCC	576	
Dy	678	CACATATTTGGCGGCTCTCTTTGGGCTCAACAGTGAACCTGATCTCTACCGAAAAAACCCTG	737	
Qy	577	GAGGGAACGAGAGATTAAGATCAACACACACACATACCCAGTTGTGTCTGCCATGCTGGC	636	
Dy	738	GATAGACACAGACAGACAGCTCAGTGTACCACTGGACCTTTTTCGCATATATTTGGC	797	
Qy	637	GCCTCTTCTTGTGATGATTTCTGGCCAAATTTCAACTGTCTGTGAGAGTGCATTC	696	
Dy	798	ACCTCTTCTTGTGATATATCTGCCCCAGTTTCAATTCAAGCAGTTCCTTCCACGAGAT	857	
Qy	697	GAAGAAGAAATGCCGTGTTCAACACTTATATGTCTGTAGACATGACAGGTGTGACAGCC	756	
Dy	858	GCCGAGACCGAGAGAGGCTCTCAATATCACTCTCTTGTGGAGGAGAGTGTCTAACACA	917	

```
QY 757 ATCTAGGGTCATCTTGGCTCAACCCCAAGAGATCAGAAAGCTTATGTGACAGT 816
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 918 GTGACAGATCATGATTTGTACACAGAGGCAAGTTGATGTGTCAATCCAGAT 977
QY 817 GCGGTGTTGAGAGAGCGGTGGCTGTGGGTACCTGTGTACCTGATCCCTTCCGTGG 876
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 978 GCCACGCTTGACAGGTGGGTGGTGTGGGACAGCTGGGAGATGATGCTCACACCTTAC 1037
QY 877 CTTCGCATGTGCTGGGTCTTGTGTGCTGGGTGATCTCCGTGGGGAGCCAGTACTG 936
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1038 GGCCCTCTCATCTGGGGGTCTTCTGCGGCAATTTCTCCACCTTAGATTTGCAATACCTA 1097
QY 937 CCGGGGTGTGTGAACCGAGTGTGGGATTCGCCACAGCTCCATCATGGGCTACAACCTTC 996
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1098 ACCGATCTTCCTGAGATCCCGCTTGGCATCCAGGACATGTGGCATTCACACCTGAC 1157
QY 997 AGCTTGCTGGGTCTGCTTGTGAGAGATCATCTAATTGTG 1035
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1158 GGCAATCTCGCATCATAGCGGCGCATGTGGGTGCTGTG 1196
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 5
US-09-864-761-1547/c
; Sequence 1547, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; PRIOR FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263,6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
```

```
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 1547
; LENGTH: 486
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AL031284.9
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 4.2
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.4
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.4
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 3.4
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.7
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2.5
US-09-864-761-1547
Query Match 11.1%; Score 139.8; DB 10; Length 486;
Best Local Similarity 92.5%; Pred. No. 4.6e-34;
Matches 147; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY 477 TATCTTCAACACAGACTACCAATGAACATGATGCATCTAGCTGTTCCAGCTATTT 536
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 458 TTATTCACAGACAGACTACCAATGAACCTGAGGCACTTACCTGTTCCAGCTATTT 399
QY 537 TGGGCTGCTGTGGGCTGGTGGCCCTGCCAAGCCTCTACCGGAGGAAGGAGATAAGA 596
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 398 TGGGCTGACTGTGGCTGGTGGCCCTGCCAAGCCTCTACCGGAGGAAGGAGATAAGA 339
QY 597 TCAGACAGCAACGATATACCAGTTGTGTGCCATCCTGGG 635
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 338 TCAGAGAGCAACGATATACCAGTTGTGTGCCATCCTGGG 300
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 6
US-09-864-761-18306/c
; Sequence 18306, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; PRIOR FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263,6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
```

```

      / PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/090,112
      / PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-22
      / PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/090,113
      / PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-22
      / NUMBER OF SEQ ID NOS: 532
      / SOFTWARE: PatentIn Ver. 2.0
      / SEQ ID NO 75
      / LENGTH: 1650
      / TYPE: DNA
      / ORGANISM: Homo sapiens
      / US-10-012-542-75

Query Match      8.7%; Score 109.6; DB 9; Length 1650;
Best Local Similarity 48.7%; Pred. Mis.3.3e-24;
Matches 327; Conservative 0; Mismatches 344; Indels 1; Gaps 1;

QY 364 TCGGCTGATCTCATGATGATGCTGCTTGTGGGGAAGCACTTGGCGCAGTGGTG 423
DB 33 TCTGCTCGCGTGGCTTTTGGGCACTTGTGGTAAGTCAAGCCCATTCAGCTGTATC 92
QY 424 ATGGTCTGTGTGAGGTGACAGCTTTAGGCAACCTGAGATGATCATCATGTAATATCTTC 483
DB 93 ATGACTTTCTTCAAGTACACCTCTTGCTGTGATGATGATGTTCACTTCCTTAACCTGCTA 152
QY 484 AACACAGCTACACATGTAACATGATGCAATCTGCTGTTTGCAGCCTATTTTGGCTG 543
DB 153 AAGTGAAGAGATGACGAGGCTCCATGACATTCACACATTTTGGGCTTACTTTGGGCTC 212
QY 544 TCTGTGGCTGTGCTGCTGCTGCCAAAGCTCTTACCAGGGAAGGAGGATAAAGATAGACA 603
DB 213 AAGTGAACCCGATCTTACCGACGCACTTGAAGCAGACAGAAAGACAGAAATTTCT 272
QY 604 GCACAGATACCCAGTTTGTCTGCAATGCTGGGCGCCCTCTTCTGTGAGTGTTCGCCA 663
DB 273 GTGTACAGTGGGACCTCTTGTGCATGATTTGGACACCTCTTCTGTGATGATGACGCC 332
QY 664 AGTTCAACTCTGCTGTGCTGTAAGATTCGAATCGAAAGAAATGCCGTTCACACC 723
DB 333 AGCTTCAACTCAGCCATATCTTACCATGGGAGCAGCCAGCACCGCCGATCAACACC 392
QY 724 TACTATGCTGTAAGCATGAGGTGTGAACCAATCTAGGGTCACTCTTGGCTCACCCC 783
DB 393 TACTCTCTCTTGGCAGCTGCGTGTCTTACCTCGGTGGCAATATC-CAGTGCCTTGCAAG 451
QY 784 CAAGGGAAGATCAGCAAGACTTATGTGACAGTGGCGGTGTGGCGAGAGGCGTGGCTTG 843
DB 452 AAGGCAAGAGCTGACATGTGTGCAATCAGAAATGCCAGCGCTGCGAGAGGGGTGGCGGTG 511
QY 844 GGTACTCGTGTACACTGATCCCTTCTCCGTGAGCTTGGCATGTGTGCGGTCTTGTGCT 903
DB 512 GGTACCGGTGTGATGATGATGTGCTATGCGTCTTAACGGTGCCTCATCATGTGGCTGTGC 571
QY 904 GGGCTGATCTCCGTGGGGGAGCCAAAGTACTGCCGGGGGTGTGTAACCAAGTGTGGGG 963
DB 572 GGCATCATCTCCACCTCGGTTTGTATACCTGACCCCATTCCTGAGATCCCGCTGAC 631
QY 964 ATTCCCAAGTCCATGATGGGCTACAACTTCAAGCTTGTGGGCTGCTGCTTGGAGAGAT 1023
DB 632 ATCCAGACATGTGGCATTTAACAATGTGCATGGCAATCTCTGGATCATATAGCGGCATC 691
QY 1024 ATCTACATTTGTG 1035
DB 692 GTGGGTGCTGTG 703

RESULT 8
US-09-864-761-20699
/ Sequence 20699, Application US/09864761
/ Patent No. US20020048763A1
/ GENERAL INFORMATION:
/ APPLICANT: Penn, Sharon G.
/ APPLICANT: Rank, David R.
/ APPLICANT: Hanzel, David K.

```

```
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCE: Aomicca-X-1
CURRENT FILING DATE: 2001-05-23
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: GB 24263.6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00670
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: US 09/774,203
PRIOR FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
SEQ ID NO 20699
LENGTH: 446
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AC007249.2
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 10
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 14
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 13
OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 12
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 13
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 13
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 12
OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 13
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 12
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 9.7
OTHER INFORMATION: NT HIT: AL161539.2, EVALUUE 3.70e+00
US-09-864-761-20699
```

```
Query Match 4.5% Score 56.6; DB 10; Length 446;
Best Local Similarity 52.3%; Pred. No. 1.1e-07;
Matches 125; Conservative 0; Mismatches 114; Indels 0; Gaps 0;
```

```
Qy 229 GTGGCTTCAACCTTCTGATGCTGGCGCTTGCTGCAAGTGGCAATCCTGCTGACGCGC 288
Dy 133 GTGGCGGTGATGTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 192
Qy 289 TTCTGAGCCAGTTCCCTTCTGGGAAGTGTGATCATCACTTTGATGATTCGGCTGGCC 348
```

```
Dy 193 GTGGTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 252
Qy 349 ACCATGAGTCTTGTGCTGCTGATCTCATGATGATGATGATGATGATGATGATGATG 408
Dy 253 GTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 312
Qy 409 GCGGAGTTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 467
Dy 313 GTGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 371
```

```
RESULT 9
US-09-864-761-3471/c
Sequence 3471, Application US/09864761
Patent No. US20020048763A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCE: Aomicca-X-1
CURRENT FILING DATE: 2001-05-23
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: GB 24263.6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00670
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: US 09/774,203
PRIOR FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
SEQ ID NO 3471
LENGTH: 1973
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AC006547.9
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 17
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 11
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 8.6
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 9
```

OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 12  
OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 11  
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 10  
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 8.4  
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 10  
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 12  
US-09-864-761-3471

Query Match 3.9%; Score 48.6; DB 10; Length 1973;  
Best Local Similarity 50.2%; Pred. No. 8e-05;  
Matches 120; Conservative 0; Mismatches 119; Indels 0; Gaps 0;

QY 229 GTGGCCCTTCAACCTTCTCATGCTGGCCGCTTGGTGTGACAGTGGCAATCTTCTGACGCGC 288  
DB 1948 GTGGTAGAGTAGAGTAGAGTAGAGTAGAGTAGAGTAGAGTAGAGTAGAGTAGAGTAGAGTAG 1889  
QY 289 TTCTTGAGCCAGTTCCTTCTGGAAGTGTCTACACACTGTTCAAGTATTCGGCTGGCC 348  
DB 1888 ATGTGTAGTG 1829  
QY 349 ACATGAGT 408  
DB 1828 GTGTGTAGTATG 1769  
QY 409 GCGCAGTTGT 467  
DB 1768 GTGATGT 1710

RESULT 10

US-09-864-761-3936  
Sequence 3936, Application US/09864761  
Patent No. US20020048763A1

GENERAL INFORMATION:

APPLICANT: Penn, Sharon G.

APPLICANT: Rank, David R.

APPLICANT: Hanzel, David K.

APPLICANT: Chen, Wensheng

TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR

TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY

FILE REFERENCE: Aemica-X-1

CURRENT APPLICATION NUMBER: US/09/864,761

PRIOR FILING DATE: 2001-05-23

PRIOR APPLICATION NUMBER: US 60/180,312

PRIOR FILING DATE: 2000-02-04

PRIOR APPLICATION NUMBER: US 60/207,456

PRIOR FILING DATE: 2000-05-26

PRIOR APPLICATION NUMBER: US 09/632,366

PRIOR FILING DATE: 2000-08-03

PRIOR APPLICATION NUMBER: GB 24263.6

PRIOR FILING DATE: 2000-10-04

PRIOR APPLICATION NUMBER: US 60/236,359

PRIOR FILING DATE: 2000-09-27

PRIOR APPLICATION NUMBER: PCT/US01/00666

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00667

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00664

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00669

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00665

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00668

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00663

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00662

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00661

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00670

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: US 60/234,687  
PRIOR FILING DATE: 2000-09-21  
PRIOR APPLICATION NUMBER: US 09/608,408  
PRIOR FILING DATE: 2000-06-30  
PRIOR APPLICATION NUMBER: US 09/774,203  
PRIOR FILING DATE: 2001-01-29  
NUMBER OF SEQ ID NOS: 49117  
SOFTWARE: Anomax Sequence Listing Engine vers. 1.1  
SEQ ID NO 3936

LENGTH: 401  
TYPE: DNA  
ORGANISM: Homo sapiens

FEATURE:  
OTHER INFORMATION: MAP TO AC007249.2

OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 10

OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 14

OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 13

OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 12

OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 13

OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 13

OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 12

OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 13

OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 12

OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 9.7

US-09-864-761-3936

Query Match 3.8%; Score 48; DB 10; Length 401;  
Best Local Similarity 50.9%; Pred. No. 5.3e-05;  
Matches 114; Conservative 0; Mismatches 110; Indels 0; Gaps 0;

QY 249 GCTGCGCTTGT 308  
DB 157 GTGCGCGT 216  
QY 309 TGGAGAGT 368  
DB 217 AGTGCGGT 276  
QY 369 GCTGATCTGATGT 428  
DB 277 GT 336  
QY 429 GCTGTGAGGT 472  
DB 337 GGT 380

RESULT 11

US-09-864-761-21778/c

Sequence 21778, Application US/09864761

Patent No. US20020048763A1

GENERAL INFORMATION:

APPLICANT: Penn, Sharon G.

APPLICANT: Rank, David R.

APPLICANT: Hanzel, David K.

APPLICANT: Chen, Wensheng

TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR

TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY

FILE REFERENCE: Aemica-X-1

CURRENT APPLICATION NUMBER: US/09/864,761

PRIOR FILING DATE: 2001-05-23

PRIOR APPLICATION NUMBER: US 60/180,312

PRIOR FILING DATE: 2000-02-04

PRIOR APPLICATION NUMBER: US 60/207,456

PRIOR FILING DATE: 2000-05-26

PRIOR APPLICATION NUMBER: US 09/632,366

PRIOR FILING DATE: 2000-08-03

PRIOR APPLICATION NUMBER: GB 24263.6

PRIOR FILING DATE: 2000-10-04

PRIOR APPLICATION NUMBER: US 60/236,359

PRIOR FILING DATE: 2000-09-27

PRIOR APPLICATION NUMBER: PCT/US01/00666

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00666

PRIOR FILING DATE: 2001-01-30

```
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annonmax Sequence Listing Engine vers. 1.1
; SEQ ID NO 21778
; LENGTH: 570
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC005691.1
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.3
; OTHER INFORMATION: NT HIT: Y11204.1, EVALUE 3.00e-04
US-09-864-761-21778
```

```
Query Match 3.8%; Score 47.4; DB 10; Length 570;
Best Local Similarity 50.2%; Pred. No. 9.9e-05;
Matches 117; Conservative 0; Mismatches 116; Indels 0; Gaps 0;
```

```
Qy 259 GGTGTCAGTGGGCAATCTCTGTGAGCGCTTCTGAGCCAGTTCCTTCTGGGAAGTG 318
Db 249 GTTAATGATGATGGGACAAATGATGATGCTCGTATGATGAGGGGTAGAGCTGTGATGTG 190
Qy 319 GTCAACACACTGTCAGTATTCGGCTGGCCACCATGATGAGTCTTTCGGTGTGATCTCA 378
Db 189 GTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 130
Qy 379 GTGATGCTGTCTTGGGGAAGGTCAACTTGGCGCAGTTGGTGTGATGCTGTGATGAG 438
Db 129 GTGATGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 70
Qy 439 GTGACGCTTTAGGCAACTGAGATGTCATCAGTAATATCTTCAACACAGA 491
Db 69 ATGAGAGTTGAGATGAGGAAGACAAATAGTATGATGATGATGATGATGATGATGATG 17
```

```
RESULT 12
US-09-864-761-5052/c
; Sequence 5052, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aecmiga-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; PRIOR FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
```

```
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annonmax Sequence Listing Engine vers. 1.1
; SEQ ID NO 5052
; LENGTH: 1976
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC005691.1
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.3
US-09-864-761-5052
```

```
Query Match 3.8%; Score 47.4; DB 10; Length 1976;
Best Local Similarity 50.2%; Pred. No. 0.00019;
Matches 117; Conservative 0; Mismatches 116; Indels 0; Gaps 0;
```

```
Qy 259 GGTGTCAGTGGGCAATCTCTGTGAGCGCTTCTGAGCCAGTTCCTTCTGGGAAGTG 318
Db 593 GTTAATGATGATGGGACAAATGATGATGCTCGTATGATGAGGGGTAGAGCTGTGATGTG 534
Qy 319 GTCAACACACTGTCAGTATTCGGCTGGCCACCATGATGAGTCTTTCGGTGTGATCTCA 378
Db 533 GTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 474
Qy 379 GTGATGCTGTCTTGGGGAAGGTCAACTTGGCGCAGTTGGTGTGATGCTGTGATGAG 438
Db 473 GTGATGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 414
Qy 439 GTGACGCTTTAGGCAACTGAGATGTCATCAGTAATATCTTCAACACAGA 491
Db 413 ATGAGAGTTGAGATGAGGAAGACAAATAGTATGATGATGATGATGATGATGATGATG 361
```

```
RESULT 13
US-09-864-761-19241
; Sequence 19241, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
```





OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 34  
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 7.1  
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 28  
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 43  
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 20  
OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 25  
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 18  
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 21  
US-09-664-761-2513

Query Match 3.7%; Score 46.6; DB 10; Length 1403;  
Best Local Similarity 51.2%; Pred. No. 0.00029;  
Matches 109; Conservative 0; Mismatches 104; Indels 0; Gaps 0;

OY 229 GTGGCTTCAACCTTTCATGCTGGCGCTGTGTGACATGGGCAATCCTGTGGACGCG 288  
DB 484 GTGGTGTGATGCTGT 543  
OY 289 TTCTGAGCCAGTTCCCTTCTGGGAGGTGTGTATCACACTTTCAATTCGCTGGCC 348  
DB 544 GTGGTGTGATGCTGT 603  
OY 349 ACCATGATGCTTTTGT 408  
DB 604 GTGGTGTGATGCTGT 663  
OY 409 GCGCAGTTGT 441  
DB 664 GTGGTGTGATGCTGT 696

## RESULT 15

US-09-664-761-20241/c

Sequence 20241, Application US/09864761

Patent No. US20020048763A1

GENERAL INFORMATION:

APPLICANT: Penn, Sharon G.

APPLICANT: Rank, David R.

APPLICANT: Hanzel, David K.

APPLICANT: Chen, Wensheng

TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR

FILE REFERENCE: Aecmlca-X-1

CURRENT FILING DATE: US/09/864,761

CURRENT FILING DATE: 2001-05-23

PRIOR APPLICATION NUMBER: US 60/180,312

PRIOR FILING DATE: 2000-02-04

PRIOR APPLICATION NUMBER: US 60/207,456

PRIOR FILING DATE: 2000-05-26

PRIOR APPLICATION NUMBER: US 09/632,366

PRIOR FILING DATE: 2000-08-03

PRIOR APPLICATION NUMBER: GB 24263.6

PRIOR FILING DATE: 2000-10-04

PRIOR APPLICATION NUMBER: US 60/236,359

PRIOR FILING DATE: 2000-09-27

PRIOR APPLICATION NUMBER: PCT/US01/00666

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00667

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00664

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00669

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00665

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00668

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00663

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00662

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00661

PRIOR FILING DATE: 2001-01-30

PRIOR FILING DATE: 2001-01-30

PRIOR FILING DATE: 2001-01-30

PRIOR FILING DATE: 2001-01-30

PRIOR FILING DATE: 2001-01-30

PRIOR FILING DATE: 2001-01-30

PRIOR FILING DATE: 2001-01-30

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00670  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: US 60/234,687  
PRIOR FILING DATE: 2000-09-21  
PRIOR APPLICATION NUMBER: US 09/608,408  
PRIOR FILING DATE: 2000-06-30  
PRIOR APPLICATION NUMBER: US 09/774,203  
PRIOR FILING DATE: 2001-01-29  
NUMBER OF SEQ ID NOS: 49117  
SOFTWARE: Annomax Sequence Listing Engine vers. 1.1  
SEQ ID NO 20241  
LENGTH: 1635  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
OTHER INFORMATION: MAP TO AC006547.9  
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 17  
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 11  
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 8.6  
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 9  
OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 12  
OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 11  
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 10  
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 8.4  
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 10  
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 12  
US-09-664-761-20241

Query Match 3.5%; Score 43.8; DB 10; Length 1635;  
Best Local Similarity 49.8%; Pred. No. 0.0024;  
Matches 111; Conservative 0; Mismatches 112; Indels 0; Gaps 0;

OY 248 TGCTGGCGCTGT 307  
DB 1142 TGCTGT 1083  
OY 308 CTGGGAGGT 367  
DB 1082 TGCTGT 1023  
OY 368 TGCTGT 427  
DB 1022 TGCTGT 963  
OY 428 TGCTGT 470  
DB 962 TGCTGT 920

Search completed: April 8, 2003, 11:26:14  
Job time : 170 secs

6 L G A A G C A G C T C T C A  
QY

RESULT 1  
 US-08-553-888A-1  
 : Sequence 1, Application US/08553888A  
 : Patent No. 5723293  
 : GENERAL INFORMATION:  
 : APPLICANT: Huang  
 : TITLE OF INVENTION: DIAGNOSTIC METHOD AND KIT FOR  
 : TITLE OF INVENTION: DETERMINING RH BLOOD GROUP GENOTYPE  
 : NUMBER OF SEQUENCES: 3  
 : CORRESPONDENCE ADDRESS:  
 : ADDRESSEE: Hoffmann & Baron, LLP  
 : STREET: 350 Jericho Turnpike  
 : City: Jericho  
 : STATE: New York  
 : COUNTRY: USA  
 : ZIP: 11753  
 : COMPUTER READABLE FORM:  
 : MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
 : COMPUTER: IBM compatible  
 : OPERATING SYSTEM: MS-DOS  
 : SOFTWARE: Wordperfect  
 : CURRENT APPLICATION DATA:  
 : APPLICATION NUMBER: US/08/553,888A  
 : FILING DATE: 11/06/95  
 : CLASSIFICATION: 435  
 : ATTORNEY/AGENT INFORMATION:  
 : NAME: O'Dea, Sean W.  
 : REGISTRATION NUMBER: 37690  
 : REFERENCE/DOCKET NUMBER: 454-5  
 : TELECOMMUNICATION INFORMATION:  
 : TELEPHONE: (516) 822-3550  
 : TELEFAX: (516) 822-3582  
 : INFORMATION FOR SEQ ID NO: 1:  
 : SEQUENCE CHARACTERISTICS:  
 : LENGTH: 1384 nucleotides  
 : TYPE: nucleic acid  
 : STRANDEDNESS: double  
 : TOPOLOGY: linear  
 : US-08-553-888A-1

Db 91 GAAGAGCTCTCATTCCTCTCTTATTTTATCCACATATGACGCTTCTTACAGAT 150  
Qy 121 CAAAGAGGCTCTGATGATCCATCAAGTTCAGCAAGATTCGACCGTATGGCGGCT 180  
Db 151 CAAAGAGGCTCTGATGATCCATCAAGTTCAGCAAGATTCGACCGTATGGCGGCT 210  
Qy 181 GGGTTGGGCTCTGATGATCCATCAAGTTCAGCAAGATTCGACCGTATGGCGGCT 240  
Db 211 GGGTTGGGCTCTGATGATCCATCAAGTTCAGCAAGATTCGACCGTATGGCGGCT 270  
Qy 241 CTCTTCATGCTGGGCTTGGTGTGATGAGGCAATCCGCTGAGAGGCTTCTGAGCCAG 300  
Db 271 CTCTTCATGCTGGGCTTGGTGTGATGAGGCAATCCGCTGAGAGGCTTCTGAGCCAG 330  
Qy 301 TTCCCTTTGGGAGAGTGTATCACTGTTTCAATTTCCGCTGGCCACCATAGAGTCT 360  
Db 331 TTCCCTTTGGGAGAGTGTATCACTGTTTCAATTTCCGCTGGCCACCATAGAGTCT 390  
Qy 361 TTGTCGGTGTATCTCAATTCATGATGCTCTTGGGAGAGTCACTTGGCGCATTTG 420  
Db 391 ATGTCGGTGTATCTCAATTCATGATGCTCTTGGGAGAGTCACTTGGCGCATTTG 450  
Qy 421 GTGATGCTGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 480  
Db 451 GTGATGCTGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 510  
Qy 481 TTCAACAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 540  
Db 511 TTCAACAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 570  
Qy 541 CTGTCGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 600  
Db 571 CTGTCGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 630  
Qy 601 ACAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660  
Db 631 ACAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 690  
Qy 661 CCAAGTTTCACTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 720  
Db 691 CCAAGTTTCACTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 750  
Qy 721 ACTTATGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 780  
Db 751 ACTTATGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 810  
Qy 781 CCCCAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 840  
Db 811 CCCCAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 870  
Qy 841 GGGGTTACCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGT 900  
Db 871 GGGGTTACCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGT 930  
Qy 901 GCTGGGCTGTATCTCCGTGGGGAGCAAGTACCTGCGGGGTGTGTAAACCGAGTGTG 960  
Db 931 GCTGGGCTGTATCTCCGTGGGGAGCAAGTACCTGCGGGGTGTGTAAACCGAGTGTG 990  
Qy 961 GGGATTTCCACAGCTCCATCATGAGGCTAACCTTCAAGTCTTCTGCTGTGCTGTG 1020  
Db 991 GGGATTTCCACAGCTCCATCATGAGGCTAACCTTCAAGTCTTCTGCTGTGCTGTG 1050  
Qy 1021 ATCATCTCATTTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGT 1080  
Db 1051 ATCATCTCATTTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGT 1110  
Qy 1081 TTCCAGGCTCTCTGATGATGAGGAACTCAAGTTCGCGCATGTATAGCTTCAAGTCT 1140  
Db 1111 TTCCAGGCTCTCTGATGATGAGGAACTCAAGTTCGCGCATGTATAGCTTCAAGTCT 1170  
Qy 1141 GGTCTCTGACAGGTTTGTCTCTAATATGTAATAATGGAAGACCTCATAGAGCTTAA 1200  
Db 1171 GGTCTCTGACAGGTTTGTCTCTAATATGTAATAATGGAAGACCTCATAGAGCTTAA 1230

Qy 1201 TATTTGATGACCAAGTTTCTGGAAGTTTCTCATTTTGGCTGTGATTTTAA 1254  
Db 1231 TATTTGATGACCAAGTTTCTGGAAGTTTCTCATTTTGGCTGTGATTTTAA 1284

RESULT 2  
US-08-553-888A-2  
Sequence 2, Application US/08553888A  
Patent No. 5723293

GENERAL INFORMATION:  
APPLICANT: Huang  
TITLE OF INVENTION: DIAGNOSTIC METHOD AND KIT FOR  
NUMBER OF INVENTIONS: DETERMINING RH BLOOD GROUP GENOTYPE  
NUMBER OF SEQUENCES: 3  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Hofmann & Baron, LLP  
STREET: 350 Jericho Turnpike  
CITY: Jericho  
STATE: New York  
COUNTRY: USA  
ZIP: 11753

COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
COMPUTER: IBM compatible  
OPERATING SYSTEM: MS-DOS  
SOFTWARE: Wordperfect  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/553, 888A  
FILING DATE: 11/06/95  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: O'Dea, Sean W.  
REGISTRATION NUMBER: 37690  
REFERENCE/DOCKET NUMBER: 454-5  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (516) 822-3550  
TELEFAX: (516) 822-3582  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1466 nucleotides  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear

US-08-553-888A-2

Query Match 94.3%; Score 1182; DB 1; Length 1466;  
Best Local Similarity 96.4%; Pred. No. 0;  
Matches 1209; Conservative 0; Mismatches 45; Indels 0; Gaps 0;

Qy 1 ATGAGCTCTAGTACCCGCGTGTGCTCCGCGCTGCTGCGCTCTGCGCCCTTAACACTG 60  
Db 45 ATGAGCTCTAGTACCCGCGTGTGCTCCGCGCTGCTGCGCTCTGCGCCCTTAACACTG 104  
Qy 61 GAAGAGCTCTCATTTCTCTCTTATTTTAAACCATATGACGCTTCTTAAAGAT 120  
Db 105 GAAGAGCTCTCATTTCTCTCTTATTTTAAACCATATGACGCTTCTTAAAGAT 164  
Qy 121 CAAAGAGGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 180  
Db 165 CAAAGAGGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 224  
Qy 181 GGGTTGGGCTCTGATGATCCATCAAGTTCAGCAAGATTCGAGGCTTCTGAGCCAG 240  
Db 225 GGGTTGGGCTCTGATGATCCATCAAGTTCGAGGCTTCTGAGGCTTCTGAGCCAG 284  
Qy 241 CTCTTCATGCTGGGCTTGGTGTGATGAGGCAATCCGCTGAGAGGCTTCTGAGCCAG 300  
Db 285 CTCTTCATGCTGGGCTTGGTGTGATGAGGCAATCCGCTGAGAGGCTTCTGAGCCAG 344  
Qy 301 TTCCCTTTGGGAGAGTGTATCACTGTTTCAATTTCCGCTGGCCACCATAGAGTCT 360  
Db 345 TTCCCTTTGGGAGAGTGTATCACTGTTTCAATTTCCGCTGGCCACCATAGAGTCT 404

Qy	361	TTGTCGGGCTGATCTCACTAGTGGATGCTCTTTGGGAAAGGCAACTTGGCGCAGTTGGT	420
Dp	405	ATGTCGGTCTGATCTCAAGCGGGTCTGTCTTGGGGAAAGGCAACTTGGCGCAGTTGGT	464
Qy	421	GTATGCTGCTGGTGGAGGTGACAGCTTTAGGCACTGAGAGTGTCAATCAGTAATATC	480
Dp	465	GTATGCTGCTGGTGGAGGTGACAGCTTTAGGCACTGAGAGTGTCAATCAGTAATATC	524
Qy	481	TTCAACACAGACTACCATATGATGACATCTAACGTTTCGAGGCTATTTTGGG	540
Dp	525	TTCAACACAGACTACCATATGATGACATCTAACGTTTCGAGGCTATTTTGGG	584
Qy	541	CTGTCTGGGCTGGTGGCTGCCAAAGCTCTTACCCGAGGGAAACGAGATTAAGATCAG	600
Dp	585	CTGACTGTGGCTGGTGGCTGCCAAAGCTCTTACCCAAAGGAAACGAGATTAAGATCAG	644
Qy	601	ACAGCAAGATACCCAGTTTGTCTGCCATGCTGGGGCGCCCTCTTCTGTGGATGTTCTGG	660
Dp	645	AGAGCAAGATACCCAGTTTGTCTGCCATGCTGGGGCGCCCTCTTCTGTGGATGTTCTGG	704
Qy	661	CCAAATTCAACTCTGCTCTCTGAGAAAGTCCAAATCGAAAGAAAGATGCGATTCAC	720
Dp	705	CCAAATGTCAACTCTCTCTCTGAGAAAGTCCAAATCGAAAGAAAGATGCGATTCAC	764
Qy	721	ACCTACTATGCTGTAGCAGTCAAGGTGTGACAGCCATCTCAGGGTCATCTTGGCTCAC	780
Dp	765	ACCTACTATGCTGTAGCAGTCAAGGTGTGACAGCCATCTCAGGGTCATCTTGGCTCAC	824
Qy	781	CCCCAAGGAAAGATACGAAACATTAATGTGACAGTGGGTGGAGAGAGGGGTGGCT	840
Dp	825	CCCCAAGGAAAGATACGATATCTTAATGTGACAGTGGGTGGAGAGAGGGGTGGCT	884
Qy	841	GTGGGTACCTCGTGTCACTGATCCCTTCTCGTGGCTTGGCCATGATGGTCTTGTG	900
Dp	885	GTGGGTACCTCGTGTCACTGATCCCTTCTCGTGGCTTGGCCATGATGGTCTTGTG	944
Qy	901	GCTGGGCTGATCTCCGTGGGGGAGCCAAAGTACTGCCCCGGGGTGTGTAAACCAAGTGGCTG	960
Dp	945	GCTGGGCTGATCTCCATGGGGGAGCCAAAGTCTGCCCCGGGGTGTGTAAACCAAGTGGCTG	1000
Qy	961	GGGATTTCCCAACAGCTCATATGGGCTACAACTTACGTTGCTGGGCTCTGTGAGAG	1020
Dp	1005	GGGATTTCAACACATCTCCGTATGACATCCCATCTTACGATTTGCTGGGCTCTGTGAGAG	1060
Qy	1021	ATCATCTACATTTGTGCTGCTGGTGTCTGAATACCGTGGAGCCGSCAAATGCGATGTTGGC	1080
Dp	1065	ATCATCTACATTTGTGCTGCTGGTGTCTGAATACCGTGGAGTGGCATATGATTTGGC	1120
Qy	1081	TTCCAGAGCTCCCTCAGGATTTGGGGAACTCAGGTTGGCCATCGGATATAGCTCAGGCT	1144
Dp	1125	TTCCAGAGCTCTCTCAGGATTTGGGGAACTCAGGTTGGCCATCGGATATAGCTCTCAGGCT	1188
Qy	1141	GGTCTCTGACAGGTTTGTCTCTTAAATTTTAAATATGAAAGCACTCATAGAGCTTAA	1200
Dp	1185	GGTCTCTGACAGGTTTGTCTCTTAAATTTTAAATATGAAAGCACTCATAGGCTTAA	1244
Qy	1201	TATTTTATGACCAAGTTTCTGGAAGTTTCTCATTTGGTGTGGATTTTAA	1254
Dp	1245	TATTTTATGACCAAGTTTCTGGAAGTTTCTCATTTGGTGTGGATTTTAA	1298

RESULT 3  
US-09-199-6378-111  
Sequence 111, Application US/091996378  
Patent No. 6355411  
GENERAL INFORMATION:  
APPLICANT: Ausubel, Frederick  
APPLICANT: Goodman, Howard M.  
APPLICANT: Ramee, Laurence G.  
APPLICANT: Mani, Jan-Niklos, Shalina  
APPLICANT: Tan, Man-Wah  
APPLICANT: Cao, Hui

```

; APPLICANT: Drenkard, Eliana
; APPLICANT: Tsongalis, John
; TITLE OF INVENTION: VIRULENCE-ASSOCIATED NUCLEIC ACID
; TITLE OF INVENTION: SEQUENCES AND USES THEREOF
; FILE REFERENCE: 00786/361002
; CURRENT APPLICATION NUMBER: US/09/199,637A
; CURRENT FILING DATE: 1998-11-25
; PRIOR APPLICATION NUMBER: 60/066,517
; PRIOR FILING DATE: 1997-11-25
; NUMBER OF SEQ ID NOS: 457
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 111
; LENGTH: 6909
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-199-637A-111

Query Match          3.3%; Score 41; DB 4; Length 6909;
Best Local Similarity 45.3%; Pred. No. 0.05;
Matches 149; Conservative 0; Mismatches 180; Indels 0; Gaps 0;

```

Qy	770	CTTTGGCTCAACCCCAAGGGAAGATAGACAACTTATGTGCAACAGTCGGTGTGGCAG	829
Db	1497	CTTCGGTAACGCCCGCTGACAGCTGGGACACCCCTGAGCTGGACGACCGAGATGTTCCG	1556
Qy	830	GAGCGCTGGCTGTGGGTACTCTGTCACACTATACCTCTTCTCCGTGGCTTGCCATGATGC	889
Db	1557	GGCTGACGCCCAAGGGTCGACACTTCTCTTAATAATTTTCCCGCGATCCGCGTACCAAG	1616
Qy	890	TGGGCTCTTGTGGCTGGGCTGATCTCCGTCGGGGGAGACCAAGTACCTGCGGGGTGTGTGA	949
Db	1617	TGATTTCCGAGGCCCTGAGCCGACCAACATCTGTGTGGCGGGCTGGCGTGGCTGACCCGAC	1676
Qy	950	ACCGAGTCTGGGGATTTCCCAACGCTCCATATGAGGCTACAACTTCAAGCTTTCGGGTC	1009
Db	1677	GCACCAACCCGCCCAAGCCCAACGCGGCTCCCGGGGACCCGCTGGAGCGGGTCGGCT	1766
Qy	1010	TGCTTGAGAGATCATCTACATTTGCTCTGCTGTGCTTGATACCTGCGAGCCGGCAATG	1069
Db	1737	CGCTGCCCGGATTCATCTCCTGTTGTATGCTGGCCGACACTCGGATCGCACCTACT	1786
Qy	1070	GCATGATTTGGCTTCCAGTCTCTCTACG	1098
Db	1797	ACATGAAGGCATCTGCCTTACAGGAC	1825

RESULT 4  
 ; Sequence 15, Application US/08476176B  
 ; Patent No. 5958708  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Hardman, No. 5958708man  
 ; APPLICANT: Kolbinger, Frank  
 ; APPLICANT: Saldanha, Jose  
 ; TITLE OF INVENTION: Reshaped monoclonal antibodies against an  
 ; TITLE OF INVENTION: Immunoglobulin Isotype  
 ; NUMBER OF SEQUENCES: 55  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: No. 5958708artis Patent Department  
 ; STREET: 59 Route 410  
 ; CITY: East Hanover  
 ; STATE: New Jersey  
 ; COUNTRY: USA  
 ; ZIP: 07936-1080  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/476,176B  
 ; FILING DATE:  
 ; CLASSIFICATION: 435

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/127,721  
FILING DATE: 27-SEPTEMBER-1993  
APPLICATION NUMBER: US 07/952,802  
FILING DATE: 25-SEPTEMBER-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: No. 5958708ak, Henry P.  
REGISTRATION NUMBER: 33,200  
REFERENCE/DOCKET NUMBER: 4-19276/A/P2/CIP  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (908) 277-5110  
TELEFAX: (908) 277-4306  
INFORMATION FOR SEQ ID NO: 15:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 467 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 22..447  
FEATURE:  
NAME/KEY: mat\_peptide  
LOCATION: 79..447  
OTHER INFORMATION: /product= "heavy chain variable"  
OTHER INFORMATION: region C21-Hay1"  
US-08-476-176B-15

Query Match 3.1%; Score 39.4; DB 2; Length 467;  
Best Local Similarity 51.4%; Pred. No. 0.031;  
Matches 91; Conservative 0; Mismatches 86; Indels 0; Gaps 0;

QY 264 GCAGTGGCAATCCTGCTGACGCGCTTCTGAGCCAGTTCCCTTCTGGGAAGTGTGAT 323  
DB 405 GTAGTCGTAAGTTGCTGCGCGCTGAAGTGGCTGAACCTGGCGAGTGAACACGGCGGTGTC 346  
QY 324 CACACTGTTCAATATTCGGCTGGCCACCATGAGTGTCTTGTGCGTGTGATCTCACTGGA 383  
DB 345 CTCGCTGTGTCAGGCTGCTCAGCTCCATGATGAGCGGCTGCTGCGGTGCTGTCGCGCTGA 286  
QY 384 TGTCTCTTGGGGAAGTCACTTGGCGCAGTTGGTGTGATGTGCTGTGGAGGT 440  
DB 285 GGTGGCTTGGCCTTGAACCTTCTGTTGATGTGGTGAAGTGCAGGCGGTGAT 229

RESULT 5  
US-08-127-721A-15/C  
Sequence 15, Application US/08127721A  
Patent No. 6066718  
GENERAL INFORMATION:  
APPLICANT: Hardman, No. 6066718man  
APPLICANT: Kolbinger, Frank  
APPLICANT: Saldanha, Jose  
TITLE OF INVENTION: Reshaped monoclonal antibodies against an  
TITLE OF INVENTION: Immunoglobulin isotype  
NUMBER OF SEQUENCES: 55  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: No. 6066718artis Patent and Trademark Department  
STREET: 59 Route 10  
CITY: East Hanover  
STATE: New Jersey  
COUNTRY: USA  
ZIP: 07936-1080  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/127,721A  
FILING DATE: 27-SEPTEMBER-1993  
CLASSIFICATION: 424

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/952,802  
FILING DATE: 25-SEPTEMBER-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: No. 6066718ak, Henry P.  
REGISTRATION NUMBER: 33,200  
REFERENCE/DOCKET NUMBER: 4-19276/A/P2/CIP  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (908) 277-5110  
TELEFAX: (908) 277-4306  
INFORMATION FOR SEQ ID NO: 15:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 467 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 22..447  
FEATURE:  
NAME/KEY: mat\_peptide  
LOCATION: 79..447  
OTHER INFORMATION: /product= "heavy chain variable"  
OTHER INFORMATION: region C21-Hay1"  
US-08-127-721A-15

Query Match 3.1%; Score 39.4; DB 3; Length 467;  
Best Local Similarity 51.4%; Pred. No. 0.031;  
Matches 91; Conservative 0; Mismatches 86; Indels 0; Gaps 0;

QY 264 GCAGTGGCAATCCTGCTGACGCGCTTCTGAGCCAGTTCCCTTCTGGGAAGTGTGAT 323  
DB 405 GTAGTCGTAAGTTGCTGCGCGCTGAAGTGGCTGAACCTGGCGAGTGAACACGGCGGTGTC 346  
QY 324 CACACTGTTCAATATTCGGCTGGCCACCATGAGTGTCTTGTGCGTGTGATCTCACTGGA 383  
DB 345 CTCGCTGTGTCAGGCTGCTCAGCTCCATGATGAGCGGCTGCTGCGGTGCTGTCGCGCTGA 286  
QY 384 TGTCTCTTGGGGAAGTCACTTGGCGCAGTTGGTGTGATGTGCTGTGGAGGT 440  
DB 285 GGTGGCTTGGCCTTGAACCTTCTGTTGATGTGGTGAAGTGCAGGCGGTGAT 229

RESULT 6  
US-08-485-246A-15/C  
Sequence 15, Application US/08485246A  
Patent No. 6072035  
GENERAL INFORMATION:  
APPLICANT: Hardman, No. 6072035man  
APPLICANT: Kolbinger, Frank  
APPLICANT: Saldanha, Jose  
TITLE OF INVENTION: Reshaped monoclonal antibodies against an  
TITLE OF INVENTION: Immunoglobulin isotype  
NUMBER OF SEQUENCES: 55  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: No. 6072035artis Patent Department  
STREET: 59 Route 10  
CITY: East Hanover  
STATE: New Jersey  
COUNTRY: USA  
ZIP: 07936-1080  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/485,246A  
FILING DATE:  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/127,721

```

1      FILING DATE: 27-SEPTEMBER-1993
2      APPLICATION NUMBER: US 07/952,802
3      FILING DATE: 25-SEPTEMBER-1992
4      ATTORNEY/AGENT INFORMATION:
5      NAME: No. 6072033ak, Henry P.
6      REGISTRATION NUMBER: 33,200
7      REFERENCE/DOCKET NUMBER: 4-19276/A/P2/CIP
8      TELECOMMUNICATION INFORMATION:
9      TELEPHONE: (908) 277-5110
10     TELEFAX: (908) 277-4306
11     INFORMATION FOR SEQ. ID NO: 15:
12     SEQUENCE CHARACTERISTICS:
13     LENGTH: 467 base pairs
14     TYPE: nucleic acid
15     STRANDEDNESS: double
16     TOPOLOGY: linear
17     MOLECULE TYPE: DNA (genomic)
18     FEATURE:
19     NAME/KEY: CDS
20     LOCATION: 22..447
21     FEATURE:
22     NAME/KEY: mat.peptide
23     LOCATION: 79..447
24     OTHER INFORMATION: /product= "heavy chain variable
25     OTHER INFORMATION: region C21-Hey1"
26     OS-08-485-246A-15

```

Query Match	3.1%;	Score 39.4;	DB 3;	Length 467;
Best Local Similarity	51.4%;	Pred. No. 0.031;		
Matches	91;	Conservative	0;	Mismatches 86; Indels 0; Gaps

  

QY	264	GCAGTGGCAATCCCTGCTGACGGCTTCTTACGCCAGTTCCTTCTGGAGAGTGGTCAT	322
Db	405	GTAGTCGTAAGTTGCTGCGCGCTGAAGTGGCTGAACCTGGCGCAGTAAGTACAGGGGGTGC	346
QY	324	CACACTGTTACGATATTGGGGCTGGGCCACCATAGTGCCTTTGTCGGTGGCTGATCTCAATGGA	388
Db	345	CTCCCTGGTCAAGGCTGCTCAAGCTCCATATGACGGGTGCTGGCGCTGGTGTGTGGGGGTGAA	286
QY	384	TGCTGTCTTGGGGAAGTCACTTGGCGCCAGTTGGTGTGATGTGTCGTGAGAGGT	440
Db	285	GATGGCCTTGGCCTTGAACCTTCTGCTTTAGTGGTGTGAAGGTGCGGGGGCTGAT	229

RESULT 7  
 US-08-476-176B-11/c  
 Sequence 11 Application US/08476176B  
 Patent No. 5958708  
 GENERAL INFORMATION:  
 APPLICANT: Hardman, No. 5958708man  
 APPLICANT: Kolbinger, Frank  
 APPLICANT: Saldanha, Jose  
 TITLE OF INVENTION: Refreshed monoclonal antibodies against an  
 NUMBER OF INVENTION: Immunoglobulin isotype  
 NUMBER OF SEQUENCES: 55  
 CORRESPONDENCE ADDRESSES:  
 ADDRESSEE: No. 5958708artis Patent Department  
 STREET: 59 Route 10  
 CITY: East Hanover  
 STATE: New Jersey  
 COUNTRY: USA  
 ZIP: 07936-1080  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08476,176B  
 FILING DATE:  
 CLASSIFICATION: 435  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: 08/127,721

```

1      FILING DATE: 27-SEPTEMBER-1993
2      APPLICATION NUMBER: US 07/952, 802
3      FILING DATE: 25-SEPTEMBER-1992
4      ATTORNEY/AGENT INFORMATION:
5          NAME: NO. 5958708ak, Henry P.
6      REGISTRATION NUMBER: 33,200
7      REFERENCE/DOCKET NUMBER: 4-19276/A/P2/CIP
8      TELECOMMUNICATION INFORMATION:
9          TELEPHONE: (908) 277-5110
10         TELEFAX: (908) 277-4306
11      INFORMATION FOR SEQ ID NO: 11:
12          SEQUENCE CHARACTERISTICS:
13              LENGTH: 468 base pairs
14              TYPE: nucleic acid
15              STRANDEDNESS: double
16              TOPOLOGY: linear
17      MOLECULE TYPE: DNA (genomic)
18      *FEATURE:
19          NAME/KEY: CDS
20          LOCATION: 22..447
21      FEATURE:
22          NAME/KEY: mat peptide
23          LOCATION: 79..447
24      OTHER INFORMATION: /product= "heavy chain variable
25      OTHER INFORMATION: region C21-H1"
26      JS-08-476-176B-11

```

	Query Match	3.1%	Score 39.4	DB 2	Length 468
	Best Local Similarity	51.4%	Pred. No. 0.031		
	Matches 91	Conservative 0	Mismatches 86	Indels 0	Gaps
QY	264	GCAATGGGCAATCCCTGTGGACGGCTTCCTGAGCCATTCCTTCTGGAGAGTGATCAT	3223		
Db	405	GTATCTCAATTGTGTGGCGCTGAAGTGTCTGAACCTGGCCAGTATACAGCGCGTCTC	3468		
QY	324	CACACTGTTCAGTAATTCGCGCTGGCCACCATGAGTCTTTGTGGTCTGATCTTCAGTGA	3833		
Db	345	CTCCGTGTGATGAGGCTCTCACTTCAGTATGAGCGGCTTTGTGTGTCTGTCGGCGGTGAA	2868		
QY	384	TGCTGTCTTGGGGAAGTCAACTTGGCGCAGTTGGCGTATATGTCGTGTGAGAT	440		
Db	285	GGTGCCCTTGCCCTTGAACTTCTGTTGTAAGTTGGTGAGTGAAGTCCCGGCGCTGAT	229		

RESULT 8  
 US-08-127-721A-11/C  
 Sequence 11, Application US/08127721A  
 Patent No. 6066718  
 GENERAL INFORMATION:  
 APPLICANT: Hardman, No. 6066718man  
 APPLICANT: Kolbinger, Frank  
 APPLICANT: Saldanha, Jose  
 TITLE OF INVENTION: Reshaped monoclonal antibodies against an  
 TITLE OF INVENTION: Immunoglobulin isotype  
 NUMBER OF SEQUENCES: 55  
 CORRESPONDENCE ADDRESSES:  
 ADDRESSEE: No. 6066718artis Patent and Trademark Department  
 STREET: 59 Route 10  
 CITY: East Hanover  
 STATE: New Jersey  
 COUNTRY: USA  
 ZIP: 07936-1080  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/127, 721A  
 FILING DATE: 27-SEPTEMBER-1993  
 CLASSIFICATION: 424  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 07/952, 802

FILING DATE: 25-SEPTEMBER-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: No. 6066718ak, Henry P.  
REGISTRATION NUMBER: 33,200  
REFERENCE/DOCKET NUMBER: 4-19276/A/P2/CIP  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (908) 277-5110  
TELEFAX: (908) 277-4306  
INFORMATION FOR SEQ ID NO: 11:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 468 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 22..447  
FEATURE:  
NAME/KEY: mat.peptide  
LOCATION: 79..447  
OTHER INFORMATION: /product= "heavy chain variable"  
OTHER INFORMATION: region C21-H1"  
US-08-127-721A-11

Query Match 3.1%; Score 39.4; DB 3; Length 468;  
Best Local Similarity 51.4%; Pred. No. 0.031;  
Matches 91; Conservative 0; Mismatches 86; Indels 0; Gaps 0;

QY 264 GCACTGGGCAATCTCTGTGACGCGCTTCCTGACCACTTCCTTGTGGAGAGTGTAT 323  
DB 405 GTAGCTGAGTGTGCTGCGCGCTGAGAGTGTGACCTGCGCGAGTGTACACGCGGTGTC 346  
QY 324 CACACTGTTCAGTATTCGGGCTGGCCACCATGATGCTTGTGCGGCTGATCTCAGTGA 383  
DB 345 CTCGTGTCTACGGCTGTCTACGCTCCATGTAGCGGTGTGTGTGTGTGTGTGTGTGTAA 286  
QY 384 TGTCTCTTGGGAGAGTCACTTGCGCCAGTTGGTGTGATGTGTGTGTGTGTGTGTGTGT 440  
DB 285 GGTGGCCTTGGCCTTGAACTTCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 229

RESULT 9  
US-08-485-246A-11/C  
Sequence 11, Application US/08485246A  
Patent No. 6072035  
GENERAL INFORMATION:  
APPLICANT: Hardman, No. 6072035man  
APPLICANT: Kolbinger, Frank  
APPLICANT: Saldanha, Jose  
TITLE OF INVENTION: Reshaped monoclonal antibodies against an  
NUMBER OF SEQUENCES: 55  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: No. 6072035artis Patent Department  
STREET: 59 Route 10  
CITY: East Hanover  
STATE: New Jersey  
COUNTRY: USA  
ZIP: 07936-1080  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/485,246A  
FILING DATE:  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/127,721  
FILING DATE: 27-SEPTEMBER-1993  
APPLICATION NUMBER: US 07/952,802

FILING DATE: 25-SEPTEMBER-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: No. 6072035ak, Henry P.  
REGISTRATION NUMBER: 33,200  
REFERENCE/DOCKET NUMBER: 4-19276/A/P2/CIP  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (908) 277-5110  
TELEFAX: (908) 277-4306  
INFORMATION FOR SEQ ID NO: 11:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 468 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 22..447  
FEATURE:  
NAME/KEY: mat.peptide  
LOCATION: 79..447  
OTHER INFORMATION: /product= "heavy chain variable"  
OTHER INFORMATION: region C21-H1"  
US-08-485-246A-11

Query Match 3.1%; Score 39.4; DB 3; Length 468;  
Best Local Similarity 51.4%; Pred. No. 0.031;  
Matches 91; Conservative 0; Mismatches 86; Indels 0; Gaps 0;

QY 264 GCACTGGGCAATCTCTGTGACGCGCTTCCTGACCACTTCCTTGTGGAGAGTGTAT 323  
DB 405 GTAGCTGAGTGTGCTGCGCGCTGAGAGTGTGACCTGCGCGAGTGTACACGCGGTGTC 346  
QY 324 CACACTGTTCAGTATTCGGGCTGGCCACCATGATGCTTGTGCGGCTGATCTCAGTGA 383  
DB 345 CTCGTGTCTACGGCTGTCTACGCTCCATGTAGCGGTGTGTGTGTGTGTGTGTGTGTAA 286  
QY 384 TGTCTCTTGGGAGAGTCACTTGCGCCAGTTGGTGTGATGTGTGTGTGTGTGTGTGTGT 440  
DB 285 GGTGGCCTTGGCCTTGAACTTCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 229

RESULT 10  
US-07-781-034-4  
Sequence 4, Application US/07781034  
Patent No. 5442050  
GENERAL INFORMATION:  
APPLICANT: Fishman, Jay A.  
TITLE OF INVENTION: Molecular Cloning of Antigens Shared By  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.  
STREET: Two Millitia Drive  
CITY: Lexington  
STATE: Massachusetts  
COUNTRY: USA  
ZIP: 02173  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/781,034  
FILING DATE: 19911018  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: Granahan, Patricia  
REGISTRATION NUMBER: 32,227  
REFERENCE/DOCKET NUMBER: MGH91-02A  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 861-6240



TELEFAX: (617) 861-9540  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1189 base pairs  
; TYPE: NUCLEIC ACID  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA to mRNA  
; HYPOTHEICAL: NO  
; ANTI-SENSE: NO  
US-07-781-034-4

Query Match 3.0%; Score 38; DB 1; Length 1189;  
Best Local Similarity 57.6%; Pred. No. 0.14;  
Matches 68; Conservative 0; Mismatches 50; Indels 0; Gaps 0;

QY 393 GGGGAAGTCACTTGGCGGCGATTGGTGATGCTGCTGTGAGGTGACGCTTTAGG 452  
DB 544 GGTGAATGTAGCCAGATGGAAGATGTGGCTGTGATGTGATGTGATGTGATGTG 603  
QY 453 CAACCTGAGATGTCATCAGTAATCTTCAACACAGACATGACATGATGATG 510  
DB 604 GATGCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 661

## RESULT 11

PCT-US92-08328-4  
; Sequence 4, Application PC/RTUS9208328  
; GENERAL INFORMATION:  
; APPLICANT: Fishman, Jay A.  
; TITLE OF INVENTION: Molecular Cloning of Antigens Shared By  
; TITLE OF INVENTION: Rat- and Human-Derived Pneumocystis Carinii  
; NUMBER OF SEQUENCES: 4  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.  
; STREET: Two Militia Drive  
; CITY: Lexington  
; STATE: Massachusetts  
; COUNTRY: USA  
; ZIP: 02173  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US92/08328  
; FILING DATE: 19920930  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/781,034  
; FILING DATE: 18-OCT-1991  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/768,166  
; FILING DATE: 30-SEP-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Granahan, Patricia  
; REGISTRATION NUMBER: 32,227  
; REFERENCE/DOCKET NUMBER: MGH91-02AA PCT  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 861-6240  
; TELEFAX: (617) 861-9540  
; TELEX: 951794  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1189 base pairs  
; TYPE: NUCLEIC ACID  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA to mRNA  
; HYPOTHEICAL: NO  
; ANTI-SENSE: NO  
PCT-US92-08328-4

Query Match 3.0%; Score 38; DB 5; Length 1189;  
Best Local Similarity 57.6%; Pred. No. 0.14;  
Matches 68; Conservative 0; Mismatches 50; Indels 0; Gaps 0;

QY 393 GGGGAAGTCACTTGGCGGCGATTGGTGATGCTGCTGTGAGGTGACGCTTTAGG 452  
DB 544 GGTGAATGTAGCCAGATGGAAGATGTGGCTGTGATGTGATGTGATGTGATGTG 603  
QY 453 CAACCTGAGATGTCATCAGTAATCTTCAACACAGACATGACATGATGATG 510  
DB 604 GATGCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 661

## RESULT 12

US-09-370-838-142  
; Sequence 142, Application US/09370838  
; Patent No. 6444425  
; GENERAL INFORMATION:  
; APPLICANT: Reed, Steven G.  
; APPLICANT: Lodes, Michael J.  
; APPLICANT: Mohamath, Roadoh  
; APPLICANT: Secrist, Heather  
; TITLE OF INVENTION: COMPOUNDS FOR THERAPY AND DIAGNOSIS OF  
; TITLE OF INVENTION: LONG CANCER AND METHODS FOR THEIR USE  
; FILE REFERENCE: 210121.475C1  
; CURRENT APPLICATION NUMBER: US/09/370,838  
; CURRENT FILING DATE: 1999-08-09  
; EARLIER APPLICATION NUMBER: US 09/285,323  
; EARLIER FILING DATE: 1999-04-02  
; NUMBER OF SEQ ID NOS: 289  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 142  
; LENGTH: 500  
; TYPE: DNA  
; ORGANISM: Homo sapien  
US-09-370-838-142

Query Match 3.0%; Score 37.8; DB 4; Length 500;  
Best Local Similarity 52.2%; Pred. No. 0.099; 77; Indels 0; Gaps 0;  
Matches 84; Conservative 0; Mismatches 77; Indels 0; Gaps 0;

QY 910 ATCTCCGTCGGGGGAGCAAGTACCTCCGGGGTGTGTAACGAGTGTGGGATTCCC 969  
DB 200 ATTAACCTGGGAAGATCAGTCCAGGACGGGGCAAGTTCTGGCGGATGCC 259  
QY 970 CACAGCTCATGATGCGCTACCACTTCACTTGTCTGCTGCTTGAAGATCATCTAC 1029  
DB 260 GAGGTCTACATCCCGCGAGCAGCATCAAGTACCTGGCGCATCCCGAGAGATCATGAC 319  
QY 1030 ATTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1070  
DB 320 ATGCTCAAGAGAGGATGTGGCCAGAGGCCGCCGCCGCCGCCGCCGCCGCCGCC 360

## RESULT 13

US-09-372-422A-17  
; Sequence 17, Application US/09372422A  
; Patent No. 6313375  
; GENERAL INFORMATION:  
; APPLICANT: Rudolf Jung  
; APPLICANT: Francois Barrieu  
; TITLE OF INVENTION: Maize Aquaporins and Uses Thereof  
; FILE REFERENCE: 0919  
; CURRENT APPLICATION NUMBER: US/09/372,422A  
; CURRENT FILING DATE: 1999-08-11  
; PRIOR APPLICATION NUMBER: US 60/098,692  
; PRIOR FILING DATE: 1998-08-31  
; NUMBER OF SEQ ID NOS: 49  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 17  
; LENGTH: 1384  
; TYPE: DNA

ORGANISM: Zea mays  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (95)...(959)  
US-09-372-422A-17

Query Match 3.0%; Score 37.2; DB 4; Length 1384;  
Best Local Similarity 56.6%; Pred. No. 0.28;  
Matches 69; Conservative 0; Mismatches 53; Indels 0; Gaps 0;

QY 1002 GCTGGTCTGCTGGAGAGATATCTACATTTGCTGCTGGCTTGTATACCGTGGAGC 1061  
DB 478 GCTGCGCTGCAAGCGCGGCTCTTCTACATGTGCAAGTCCCGGCGCATCTGCGG 537  
QY 1062 CGGCAATGCAATGATTTGGCTTCAGATCTCCTCAGCATTTGGGAACTCAGCTTGCCAT 1121  
DB 538 CGCGGCGCTGTCAGAGGCGCTTCAGAGAGGCGCTTACATGGCGCGGCGCGCCCA 597  
QY 1122 CG 1123  
DB 598 CG 599

RESULT 14  
US-09-668-680-5  
Sequence 5, Application US/09668680  
Patent No. 6436703  
GENERAL INFORMATION:  
APPLICANT: Tang, Y. Tom  
APPLICANT: Liu, Chenghua  
APPLICANT: Zhou, Ping  
APPLICANT: Asundi, Vinod  
APPLICANT: Zhang, Jie  
APPLICANT: Wang, Jian-Rui  
APPLICANT: Xue, Aidong J.  
APPLICANT: Xu, Chongjun  
APPLICANT: Dmanac, Radoje T.  
TITLE OF INVENTION: No. 6436703el Nucleic Acids and  
FILE REFERENCE: 790CIP2A  
CURRENT APPLICATION NUMBER: US/09/668,680  
PRIOR FILING DATE: 2000-09-22  
PRIOR APPLICATION NUMBER: 09/649,167  
PRIOR FILING DATE: 2000-08-23  
PRIOR APPLICATION NUMBER: 09/540,217  
NUMBER OF SEQ ID NOS: 13  
SOFTWARE: pc\_fl\_genes Version 2.0  
SEQ ID NO 5  
LENGTH: 897  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (1)...(897)  
US-09-668-680-5

Query Match 2.9%; Score 36.6; DB 4; Length 897;  
Best Local Similarity 50.3%; Pred. No. 0.32;  
Matches 90; Conservative 0; Mismatches 89; Indels 0; Gaps 0;

QY 374 TCTCAGTATCTCTCTTGGGGAAGTCACTTGGCGGAGTTGGTGTGATGCTG 433  
DB 293 TCTCAGAGATCAAGACTGCGAGCCCGTCTCTTGGGCTGTTCTCTCCATGTACTGA 352  
QY 434 TGAAGTGAACCTTTAGGCAACTGAGAGTGTATCAATAATCTTCAACAGACT 493  
DB 353 TCACGCTGCTGGGAACCTGCTATCATCTGCGCGTGTGATCTCCACCTCCACA 412  
QY 494 ACCAATGAACATGATGACATCTACGTGTGCAAGCCATTTTGGGCTGTGTGGCC 552  
DB 413 CCCCATTACTTCTCTCTCAACTGTCTCTGGCTGACATGGAATCACTCACTCTGCC 471

RESULT 15  
US-08-332-766A-1  
Sequence 1, Application US/08332766A  
Patent No. 5843647  
GENERAL INFORMATION:  
APPLICANT: JEFFREYS, Alec J.  
APPLICANT: ARMOUR, John  
TITLE OF INVENTION: SIMPLE TANDEM REPEATS  
NUMBER OF SEQUENCES: 125  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: CUSHMAN DABRY & CUSHMAN, L.L.P.  
STREET: 1100 New York Avenue, N.W.  
CITY: Washington  
STATE: D.C.  
COUNTRY: U.S.A.  
ZIP: 20005-3918

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/332,766A  
FILING DATE: 01-NOV-1994  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: GB 9326052.9  
FILING DATE: 21-DEC-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: BIRD, Donald J.  
REGISTRATION NUMBER: 25,323  
REFERENCE/DOCKET NUMBER: 217211/M94/0434/GB  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 861-3000  
TELEFAX: (202) 822-0944  
TELEX: 6714627 CUSH  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 377 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
US-08-332-766A-1

Query Match 2.9%; Score 36.4; DB 2; Length 377;  
Best Local Similarity 59.8%; Pred. No. 0.22;  
Matches 61; Conservative 0; Mismatches 41; Indels 0; Gaps 0;

QY 366 GATGCTGATCTCAGTATGCTGCTTGGGAGGTCACTTGGCGCAGTTGTTGTGAT 425  
DB 82 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 141  
QY 426 GATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 467  
DB 142 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 183

Search completed: April 8, 2003, 05:44:34  
Job time : 79 secs

GenCore version 5.1.4.p5\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 7, 2003, 18:32:06 ; Search time 341 Seconds

(without alignments)  
8281.542 Million cell updates/sec

Title: US-09-600-714-41

Perfect score: 1254  
Sequence: 1 acgagctcaagtraccgcg.....attgctgtcgtgattttaa 1254

Scoring table: IDENTITY\_NUC  
Gapop 10.0, Gapext 1.0

Searched: 2185239 seqs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database :

```
*_N_Geneseq_101002.*
1: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.*
2: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.*
3: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.*
4: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.*
5: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.*
6: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.*
7: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.*
8: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.*
9: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.*
10: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.*
11: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.*
12: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.*
13: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.*
14: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.*
15: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.*
16: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.*
17: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.*
18: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.*
19: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.*
20: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.*
21: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.*
22: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.*
23: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.*
24: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1254	100.0	1254	20	AAK86522
2	1177	14.1	1805	24	ABK12333
3	165.4	13.2	1853	21	AAK64286
4	165.4	13.2	1949	22	AAK25763
5	165.4	13.2	1952	24	ABK49215
6	165.4	13.2	1970	21	AAK76344
7	159.4	12.7	1679	24	ABK12334
8	152.4	12.2	3706	22	AAH18728
9	150.2	12.0	2098	24	ABK49216

C	10	139.8	11.1	486	22	ABK42886	Human breast cell
C	11	139.8	11.1	486	22	ABK53307	Human foetal liver
C	12	139.8	11.1	486	22	ABK23081	Probe #1547 for ge
C	13	139.8	11.1	486	22	AAK01568	Human brain expres
C	14	139.8	11.1	486	22	AAK27012	Human bone marrow
C	15	139.8	11.1	486	22	AAK11615	Probe #1548 for ge
C	16	139.8	11.1	486	22	AAK13211	Probe #1597 used t
C	17	139.8	11.1	486	22	AAK01542	Probe #1531 used t
C	18	139.8	11.1	486	22	ABK01595	Human genome-deriv
C	19	120	9.6	1571	24	ABK09045	Human polynucleoti
C	20	113.8	9.1	1940	23	ABK02133	Drosophila melanog
C	21	111.8	8.9	123	22	ABK48015	Human breast cell
C	22	111.8	8.9	123	22	ABK65900	Human foetal liver
C	23	111.8	8.9	123	22	ABK32986	Probe #11452 for g
C	24	111.8	8.9	123	22	ABK14315	Human brain expres
C	25	111.8	8.9	123	22	AAK40045	Human bone marrow
C	26	111.8	8.9	123	22	AAK20829	Probe #10762 for g
C	27	111.8	8.9	123	22	AAK46064	Probe #14750 used
C	28	111.8	8.9	123	22	AAK06540	Probe #6531 used t
C	29	111.8	8.9	123	22	ABK14127	Human genome-deriv
C	30	109.6	8.7	1650	21	AAK297083	Human secreted pro
C	31	108.6	8.7	484	22	AAK190968	Human polynucleoti
C	32	108.6	8.7	16918	24	ABK33617	Human immune syste
C	33	101.2	8.1	1765	24	AAK28951	Human MOLT CDNA.
C	34	93.2	7.4	1222	22	AAK98147	Human EST-derived
C	35	82.2	6.6	16918	22	ABK33616	Human immune syste
C	36	79.2	6.3	730	22	ABK31425	Human secreted pro
C	37	79.2	6.3	730	24	ABK90632	Human polynucleoti
C	38	61.2	4.9	198	24	ABK75889	Human glycoprotein
C	39	57	4.5	449	22	AAK93284	CDNA encoding SRT
C	40	56.6	4.5	446	22	ABK50433	Human breast cell
C	41	56.6	4.5	446	22	ABK68382	Human foetal liver
C	42	56.6	4.5	446	22	ABK35379	Probe #13845 for g
C	43	56.6	4.5	446	22	AAK16757	Human brain expres
C	44	56.6	4.5	446	22	AAK42529	Human bone marrow
C	45	56.6	4.5	446	22	AAK123276	Probe #13209 for g

#### ALIGNMENTS

RESULT 1	
AAK86522	AAK86522 standard; cDNA; 1254 BP.
ID	AAK86522;
XX	
AC	AAK86522;
XX	
DT	04-OCT-1999 (first entry)
XX	
DE	cDNA sequence of the prevalent allele of the Rhd gene.
XX	
KW	Allele; Rheus D antigen; Rhd; weak D phenotype; blood transfusion; ss.
XX	
OS	Homo sapiens.
XX	
FH	Key
FT	CD5
FT	Location/Qualifiers
	1..1254
	/*tag= a
XX	
XX	WO9937763-A2.
XX	
XX	29-JUL-1999.
PD	
XX	
PF	18-DEC-1998; 98WO-EP08319.
XX	
PR	23-JAN-1998; 98EP-0101203.
XX	
PA	(DRKB-) DRK BLUTSPENDDIENST BADEN WUERTTEMBERG.
XX	
PI	Flegel WA, Wagner FF;
XX	
XX	WPI; 1999-469127/39.
DR	P-PSDB; AAY24056.

XX Nucleic acid sequences correlated with Rhesus weak D phenotype.  
PT useful for screening blood from donors and recipients for  
PT transfusion methods

PS Disclosure; Fig 2; 64pp; English.

The present sequence represents the prevalent allele of the Rhensu D (Rhd) antigen gene. The specification describes a Rhd contributing to or indicative of the weak D phenotype, where the Rhd polynucleotide carries at least one missense mutation as compared to the wild-type Rhd, in its transmembrane and/or intracellular regions, especially in amino acid positions 2-16, 114-149, 179-225 or/and 267-397, with the proviso that the D antigen does not carry a single missense mutation leading to a F223I or T231I substitution. The probes and antibodies are useful in the methods for detection of weak D phenotypes. Red blood cells, from probands, are useful for the assessment of the affinity, avidity and/or reactivity of monoclonal anti-D antibodies, polyclonal anti-D antisera or of anti-globulin or anti-human-globulin antisera. Detecting the presence of the Rhd associated with weak D phenotype is useful for determining that a patient in need of a blood transfusion is to be transfused with Rhd negative blood from a donor. Alternatively, testing for weak D phenotype Rhd in the blood of a donor is useful for determining whether the donor blood should be excluded for transfusion to patients having wild type Rhd or weak D types, other than that of the donor weak D type.

**SQ** Sequence 1254 BP; 241 A; 322 C; 351 G; 340 T; 0 other;

Query Match	100.0%	Score 1254	DB 20	Length 1254
Best Local Similarity	100.0%	Pred. Nc	0	
Matches 1254	0	Mismatches	0	Gaps 0

Qy	1	ATGAGCTTAAGTACCCCGGCTGTGTCGGGCGTGCCTGCCCCCTGAGGCCCTAAACATG	60
Dp	1	ATGAGCTTAAGTACCCCGGCTGTGTCGGGCGTGCCTGCCCCCTGAGGCCCTAAACATG	60
Qy	61	GAGAGCACTCTCATTTCTCTCTTATTTTAAACCATATGAGCCTTCTTATAGAGAT	120
Dp	61	GAGAGCACTCTCATTTCTCTCTTATTTTAAACCATATGAGCCTTCTTATAGAGAT	120
Qy	121	CAAAAAGGGGCTCGTGGCATCTCATCAAGTTGGCCAAATCTGACCCGTGATGGGGCCATT	180
Dp	121	CAAAAAGGGGCTCGTGGCATCTCATCAAGTTGGCCAAATCTGACCCGTGATGGGGCCATT	180
Qy	181	GAGCTTGGGCTTCTCACCTCCAGTTTCCGAGACACACAGCTGGACAGTGTGGCTTCAAC	240
Dp	181	GAGCTTGGGCTTCTCACCTCCAGTTTCCGAGACACACAGCTGGACAGTGTGGCTTCAAC	240
Qy	241	CTTTTCATGCTGGCGCTTGTTGTGCAGTGGGCAATCTGTGACGGCTTCTGAGCCAG	300
Dp	241	CTTTTCATGCTGGCGCTTGTTGTGCAGTGGGCAATCTGTGACGGCTTCTGAGCCAG	300
Qy	301	TTCCCTTTGGGAAAGTGTGATCACACTGTTCAGTATTTGGCTGGGACCATGAGTGT	360
Dp	301	TTCCCTTTGGGAAAGTGTGATCACACTGTTCAGTATTTGGCTGGGACCATGAGTGT	360
Qy	361	TTTGTGGGTGATCTCAGTGGATGCTCTCTTGGGAAAGCTCACTTGGCGCAGTTGGTG	420
Dp	361	TTTGTGGGTGATCTCAGTGGATGCTCTCTTGGGAAAGCTCACTTGGCGCAGTTGGTG	420
Qy	421	GTGATGTGCTGTGGAGGTGACAGCTTTAGGCAACCTGAGGATGTCTATCAGTAATATC	480
Dp	421	GTGATGTGCTGTGGAGGTGACAGCTTTAGGCAACCTGAGGATGTCTATCAGTAATATC	480
Qy	481	TTCAACACAGACTACACATGAAATGATGACACTTACGTGTTCGACGCTATTTTGGG	540
Dp	481	TTCAACACAGACTACACATGAAATGATGACACTTACGTGTTCGACGCTATTTTGGG	540
Qy	541	CTGTGTGTGGCTGGTGTGCTGCCAAAGCTCTACCCGAGGGGAAACGAGAGATAAAGATCAG	600
Dp	541	CTGTGTGTGGCTGGTGTGCTGCCAAAGCTCTACCCGAGGGGAAACGAGAGATAAAGATCAG	600

Qy	601	ACAGGAAGATACCAGATTGTCTGGCAATGCTGGGAGCCCTCTTCTTGTAAGTTCTGG	660
Db	601	ACAGGAAGATACCAGATTGTCTGGCAATGCTGGGAGCCCTCTTCTTGTAAGTTCTGG	660
Qy	661	CCAAAGTTTCAACTCTGCTCTGCTGAGAAATCCAAATGGAAGAAATGCGTGTTCAAC	720
Db	661	CCAAAGTTTCAACTCTGCTCTGCTGAGAAATCCAAATGGAAGAAATGCGTGTTCAAC	720
Qy	721	ACCTACTATGCTGTATGAGATCAGGTGTGTACAGCCATCTCAGGGTCAATCCTTGAGCTCAC	780
Db	721	ACCTACTATGCTGTATGAGATCAGGTGTGTACAGCCATCTCAGGGTCAATCCTTGAGCTCAC	780
Qy	781	CCCCAAGGAAGATCAGCAAGACTTATATGTGACACAGTCGGGTGTGTGACGAGAGGCGTGGCT	840
Db	781	CCCCAAGGAAGATCAGCAAGACTTATATGTGACAGTCGGGTGTGTGACGAGAGGCGTGGCT	840
Qy	841	GTGGGTACCTTGCTGTACCTGATCCCTTCTCCGCTGGCTTGCCATGGTGTGCTGGGCTCTTGTG	900
Db	841	GTGGGTACCTTGCTGTACCTGATCCCTTCTCCGCTGGCTTGCCATGGTGTGCTGGGCTCTTGTG	900
Qy	901	GCTGGGGAGTACTCCGTGCGGGGGAGCCAAAGTCTGCGGGGGTGTGTAAACGAGTCTG	960
Db	901	GCTGGGGAGTACTCCGTGCGGGGGAGCCAAAGTCTGCGGGGGTGTGTAAACGAGTCTG	960
Qy	961	GGGATTTCCCAACAGCTCCATCATGGGGTACAATTCTACGTTGCTGGGTCTGCTTGGAGAG	1020
Db	961	GGGATTTCCCAACAGCTCCATCATGGGGTACAATTCTACGTTGCTGGGTCTGCTTGGAGAG	1020
Qy	1021	ATCATCTACATTTGTGCTGTGCTGCTTGTATACCGTCCGAGACCCGGCAATGGCAATGATTTGGC	1080
Db	1021	ATCATCTACATTTGTGCTGTGCTGCTTGTATACCGTCCGAGACCCGGCAATGGCAATGATTTGGC	1080
Qy	1081	TTCCAGGTCCTCTCAGCAATGGGGAACTCAGCTTGGCCATCTCGATAGTCTTCAACGCTCT	1144
Db	1081	TTCCAGGTCCTCTCAGCAATGGGGAACTCAGCTTGGCCATCTCGATAGTCTTCAACGCTCT	1144
Qy	1141	GGTCTCTGACAGGTTTGCTCTTAATCTTAAATATAGAAAGCACTCATAGAGCTAAA	1200
Db	1141	GGTCTCTGACAGGTTTGCTCTTAATCTTAAATATAGAAAGCACTCATAGAGCTAAA	1200
Qy	1201	TATTTTATGACCAAGTTTCTGGAAGTTTCTCATTTGGCTGTGTGATTTTAA	1254
Db	1201	TATTTTATGACCAAGTTTCTGGAAGTTTCTCATTTGGCTGTGTGATTTTAA	1254

RESULT 2  
ABK12333  
ID ABK12333 standard; cDNA; 1805 BP.

AC ABK12333;

DT 05-JUN-2002 (first entry)

DE cDNA encoding human nonerythroid Rh glycoprotein RhBG

KW Human; nonerythroid Rh glycoprotein; Rhbg; RhBG; antibody; immunogen;

ion transporter; chromosome 1q21\_3; gene; ss.

OS Homo sapiens.

FH	Key	Location/Qualifiers
223		301.115

FT / \*tag= a

FT	misc_feature	1248..1377
----	--------------	------------

```

/note= "Encodes C-tail"

```

```
/*tag= c
```

FT	/note= "Atypical polyadenylation site"
----	--

PN MO200216396-A2.  
 XX  
 PD 28-FEB-2002.  
 XX  
 PF 17-AUG-2001; 2001MO-US25881.  
 XX  
 PR 21-AUG-2000; 2000US-226767P.  
 XX  
 PA (NYBL-) NEW YORK BLOOD CENT INC.  
 XX  
 PI Huang C, Liu Z;  
 XX  
 DR WPI; 2002-280905/32.  
 DR P-PSDB; AAU78091.  
 XX  
 PT Homologues of mouse or human nonerythroid Rh glycoproteins, Rhbg or Rhbg,  
 PT respectively, useful in the production of antibodies which are useful  
 PT for detecting Rhbg or Rhbg glycoproteins in a sample -  
 XX  
 PS Claim 1; Fig 1a; 59p; English.  
 XX

CC The present invention relates to a new protein or peptide comprising an  
 CC amino acid sequence having at least 60% identity to a sequence comprising  
 CC 455 (mouse nonerythroid Rh glycoprotein homologue (Rhbg)) or 458 (human  
 CC nonerythroid Rh glycoprotein homologue (Rhbg)) amino acids, fully defined  
 CC in the specification. The antibody of the invention is useful for  
 CC detecting an Rhbg or an Rhbg glycoprotein in a sample, by contacting the  
 CC sample with antibody under conditions suitable for binding, assessing the  
 CC specific binding to the antibody, and thus detecting the presence of an  
 CC epitope of Rhbg or Rhbg in the sample. The nucleic acids of the invention  
 CC are useful as probes for detecting transporter genes and particularly Rh  
 CC type B transporter genes including e.g. NH<sup>4+</sup> ion transporters. The  
 CC present nucleic acid sequence is that of the human Rhbg gene located on  
 CC chromosome 1q21.3. This sequence encodes the human Rhbg protein of the  
 CC invention. Rhbg is a polypeptide transporter-type protein.  
 CC

Sequence 1805 BP; 351 A; 550 C; 515 G; 389 T; 0 other;

Query Match 14.1%; Score 177; DB 24; Length 1805;  
 Best Local Similarity 49.5%; Pred. No. 1e-40;  
 Matches 456; Conservative 0; Mismatches 465; Indels 0; Gaps 0;

QY 141 CTATCAAGTTGGCCAAAGATCTGACCGTGTGAGCGGCCCATTTGGCTTCTTCACCTC 200  
 DB 218 CTACCAAGCTTCCAGAGCGTGCATGCGATGCTGTGGGCTTTGGCTTCTTCATGCT 277  
 QY 201 GAGTTTCGGGAGACAGCTGGAGAGTGTGCGCTTCAACTTTCATGCTGGCGCTTGG 260  
 DB 278 CTTCCTGAGCGTTACCGCTTTCAGCAGGCTGGGCTTCACTTCTCTGCGCGCTTGGC 337  
 QY 261 TGTGCAAGTGGGCAATCCCTGCTGACCGGCTTCTGAGCCAGTTCCTTGTGGAAGGTGT 320  
 DB 338 CTTGCAAGTGGGCAATCCCTGCTGACCGGCTTCTGAGCCAGTTCCTTGTGGAAGGTGT 397  
 QY 321 CATCACTGTTCAATTCCTGCTGACCGGCTTCTGAGCCAGTTCCTTGTGGAAGGTGT 380  
 DB 398 TGTGGGCGTGGAGAGCATGATGCTGACTTTTGTGCGGGGCGTGCATCTCTCTT 457  
 QY 381 GGAATGCTGTCTTTGGGAAAGTTCATTGGCGGAGTTGGTGTGATGCTGTGGAGAGT 440  
 DB 458 TGTGCGCGTCTCTGGGCAAGACCGGCGCTTACCCAGCTGCTGCTCAAGGCGCTGTGAGAGT 517  
 QY 441 GACAGCTTTAGGCAACCTGAGGATGTGATCAGTAATATCTTCAACAGACTACCAT 500  
 DB 518 GGTGCTGTTGGCATCAATGATGTTGTGCTTCTTCTTGGGGGTGAAGATGCCGG 577  
 QY 501 GAACATGATGACATCTAGTGTTCGACGCTATTTTGGGCTGTCTGTGGCTGTGCT 560  
 DB 578 AGGCTTCATGACTATCCACACCTTTGTGTGCTTCACTTGGGGCGTCTTGTGGCGTTCT 637  
 QY 561 GCCAAGCTCTTACCCGAGGGAACGAGAGATTAAGATCAGACGACGATACCATCTT 620  
 DB 638 GTACAGGCCCAAGCTGAGGAAGACACCGCCAGGGCTCTCGTCTACATTCAGACT 697

QY 621 GTCTGCATGCTGGGCGCCCTCTTCTTGTGATGTTCTGGCCAAAGTTCAACTTGTCT 680  
 DB 698 CTTCGCATGATGGGACCATCTTCTGTGATCTTCTGTGCTTCAATGCTGCACT 757  
 QY 681 GCTGAGAAATCCATGCAAAAGAAATGCGGTGTTCAACACTTACTTCTGTAGAGT 740  
 DB 758 CACAGCGCTGGGGGCTGGGAGCATGGAAGCGCCCTCAACATATCACTCCCTGGCTGC 817  
 QY 741 CAGCGTGTGACAGCATCTCAGGTCATCTTGGCTCAACCCCAAGGAAGATCAGCA 800  
 DB 818 CAGCACCTTGGCACCTTCTTCTTGTGAGCCCTTGTAGGGAAGATGGAGGCTTGACAT 877  
 QY 801 GACTTATGTGACAGTGCCTGTGTGGAGAGCGTGTGCTGTGAGTAACCTGTGTCACT 860  
 DB 878 GGTTCATCATCAAAATGACAGCGCTGTGAGAGGGTGTGTGGGACCTCAAGTAAAT 937  
 QY 861 GATCCCTTCTCCGCTGTGCTTCCATGCTGTGCTTGTGGCTGTGCTGTGCTGTGCT 920  
 DB 938 GATGCTGACACCTTGGGGCTGTGCGACGCTGCTTCTTGTGGAGCTGTCTCAAGCT 997  
 QY 921 GGGAGCCCAATACCTGCGGGGCTGTGTACCGAGTCTGGGATTCGCCACAGCTCCAT 980  
 DB 998 GGGGTACAGATTTCTTACGCGCCATCTTGAATCAAAATTAAGTCAAGACATGTG 1057  
 QY 981 CATGGCTTACAACTTCACTTGTGTGGCTGTGTGAGAGATCATCTTGTGTGCT 1040  
 DB 1058 AGTTCACAACTCCATGAGATGCGGGGCTGTGGGGGCTCTGTGGGGCTCTGTGGG 1117  
 QY 1041 GGTGCTTGAATACGCTGGAGC 1061  
 DB 1118 TGGACTTGGCACCATCATGAGC 1138

RESULT 3  
 AAC64286  
 ID AAC64286 standard; cDNA; 1853 BP.  
 XX  
 AC AAC64286;  
 XX

DT 23-FEB-2001 (first entry)

DE Human membrane-associated protein HUMAP-13 cDNA.

KW Human membrane-associated protein; HUMAP; transgenic organism;  
 KW drug screening; cell signalling modulator; agonist; antagonist;  
 KW cell differentiation modulator; cell proliferation modulator;  
 KW cell proliferative disorder; cancer; cell differentiation disorder;  
 KW developmental disorder; cell signalling disorders; endocrine disorder;  
 KW hyperplasticity; hypothyroidism; hyperparathyroidism; infection;  
 KW pancreatic disorder; diabetes mellitus; immunological disorder;  
 KW hereditary neuropathy; gonadal steroid hormone associated disorder;  
 KW infertility; ss.  
 KW

OS Homo sapiens.

PN WO200065054-A2.

PD 02-NOV-2000.

PF 20-APR-2000; 2000MO-US10884.

PR 23-APR-1999; 99US-0130694.

PR 23-JUN-1999; 99US-0140580.

PA (INCY-) INCYTE GENOMICS INC.

PI Hillman JL, Bandman O, Tang YL, Lal P, Yue H, Reddy R, Azimzai Y,

DR Baughn MR;

DR WPI; 2000-687346/67.

DR P-PSDB; AAB29656.

PT Human membrane-associated protein, useful for diagnosis and treatment  
 PT of cell signaling, cell differentiation and cell proliferation  
 PT disorders such as cancer, and for identifying agonists and antagonists  
 PS  
 PS Claim 4, Page 96; 99pp; English.

XX The invention relates to 17 human membrane-associated proteins,  
 CC HUMAP-1 to HUMAP-17 (AAB29644-B29660) and the cDNAs encoding them  
 CC (AAC64274-C64290). The invention also relates to expression constructs,  
 CC host cells and transgenic organisms comprising a HUMAP nucleic acid  
 CC sequence; the recombinant preparation of a HUMAP; methods of screening  
 CC compounds for their ability to modulate HUMAP activity or expression;  
 CC and pharmaceutical compositions comprising a HUMAP protein, a HUMAP  
 CC agonist or HUMAP antagonist. The HUMAPs acts as modulators of cell  
 CC signalling, differentiation and proliferation. A HUMAP is useful for  
 CC screening a compound for effectiveness as an agonist or antagonist of  
 CC HUMAP activity. The protein, or the identified agonist or antagonist is  
 CC useful for treating a disease or condition associated with decreased or  
 CC increased expression of functional HUMAP. A HUMAP nucleic acid is useful  
 CC for screening a compound for its ability to alter expression of that  
 CC particular HUMAP gene. A wide variety of disease may be treated using  
 CC compositions of the invention. These diseases include cell proliferative  
 CC disorders (e.g., actinic Keratosis, arteriosclerosis); cancer (e.g.,  
 CC breast, bladder, bone marrow, brain and uterine cancer); cell  
 CC differentiation disorders, in particular developmental disorders (e.g.,  
 CC renal tubular acidosis, anaemia, Cushing's syndrome, achondroplasia,  
 CC epilepsy, and muscular dystrophy); cell signalling disorders, in  
 CC particular endocrine disorders such as hypohalams and pituitary  
 CC disorders resulting from lesions such as thrombosis; disorders  
 CC associated with hyperparathyroidism (e.g., acromegaly); disorders associated  
 CC with hypohyroidism (e.g., goitre); hyperparathyroidism; pancreatic  
 CC disorders such as type I or type II diabetes mellitus; infections;  
 CC immunological disorders; hereditary neuropathies (e.g., infections;  
 CC neurofibromatosis); and disorders associated with gonadal steroid  
 CC hormones (e.g., infertility, endometriosis, polycystic ovary syndrome,  
 CC osteoporosis, Leydig cell deficiency and gynecomastia). Antibodies which  
 CC specifically bind HUMAP may be used for the diagnosis of disorders  
 CC associated with the expression of HUMAP, or in assays to monitor patients  
 CC being treated with HUMAP or agonists, antagonists or inhibitors of HUMAP.  
 CC The present sequence represents a HUMAP cDNA of the invention.  
 XX  
 XX Sequence 1853 BP; 367 A; 563 C; 508 G; 415 T; 0 other;

Query Match 13.2%; Score 165.4; DB 21; Length 1853;  
 Best Local Similarity 49.1%; Pred. No. 2.4e-37;  
 Matches 439; Conservative 0; Mismatches 456; Indels 0; Gaps 0;

QY 141 CTATCAAGTTGGCCAGATCTGACCGTGAATGCGGCATTTGGCTTGGCTTCTTCACTC 200  
 DB |||||  
 DB 207 CTACCAAGCTTCCAGACGCTGACGTATGTCTTGGGGCTTCGGCTTCTCATGAC 266  
 QY 201 GAGTTTCCGAGACACAGCTGAGAGTGGCTTCACTTTCATGCTGGCGCTTGG 260  
 DB |||||  
 DB 267 TTTCTGACGGCTTACGGCTTCAAGCGCGTGGCTTCACTTCTTGGGACGCTTGG 326  
 QY 261 TGTGAGTGGGCAATCTGCTGAGCGGCTTCTGAGCCAGTTCCCTTGGGAAAGTGT 320  
 DB |||||  
 DB 327 CATCAGTGGGGCTGCTTCATGACAGGCTGTTCCATTTTCAAGACCGTTACATGCT 386  
 QY 321 CATCACTGTTTCACTATTCGCTGGCCACATGAGTGTGTTGCTGCTGATCTCACT 380  
 DB |||||  
 DB 387 CGTGGCGGTGAGAACCTTCATCAGCTGACTTTCGCGGGCTCTGCTGTGGTGGCTT 446  
 QY 381 GGATGCTGTTGGGGAAGTCAACTGGCGAGTTGGTGGTGGTGGTGGTGGTGGTGG 440  
 DB |||||  
 DB 447 TGGGCAAGTTCTGGGTAAAGTCAAGCCCATTCAGCTGCTCATCAGACTTTCTTCAAGT 506  
 QY 441 GACAGCTTTAGGCACTGAGAGTGTATCAGTAAATCTTCAACACAGATACCAT 500  
 DB |||||  
 DB 507 GACCTCTTCTGCTGATGATGATGATCTTCTTAACTGCTTAAAGTGAAGAGATGACAG 566  
 QY 501 GAACATGATGACATCTACGTGTTGGACGCTATTTTGGGCTGTGTGGCTGTGGCT 560

DB 567 AGGCTTCATGACCATTCACACATTTGGCGGCTACTTTGGGCTCAACATGACCCGATCT 626  
 QY 561 GCCAAGCCTCTTACCGAGGGAACGAGGATTAAGATCAGACAGCAAGATACCACTTT 620  
 DB |||||  
 DB 627 CTACCGACGCACTTAAGACAGACAGAGACAGAAATTCTGTATCAAGTGGACCT 686  
 QY 621 GTCTCCATGCTGGCGCCCTCTTCTGTGATGATTTGTGGCCAAATTCACTGTCTCT 680  
 DB |||||  
 DB 687 CTTCGATGATTTGGACACCTCTTCTGTGATGATTTGGCCGCTTCAACTCAGCAT 746  
 QY 681 GCTAGAAATTCATTCGAAAGAAATGCGGTCTTCAACCTACTATGTCTGAGAGT 740  
 DB |||||  
 DB 747 ATCTTACCATGGGAGACAGACAGACCGAGCGCCATCAACCTACCTCTTGGGAGC 806  
 QY 741 CAGGCTGTGACAGCATCTCAGGGTATCTGCTTGTGCTCACCCCAAGGAAATCAGAA 800  
 DB |||||  
 DB 807 CTGCTGCTTACCTCGGTGGGAATATCAAGTCCCTGACAGAGGCAAGCTTGACAT 866  
 QY 801 GACTTATGTGACAGTGGCTGTTGGCAGAGGCGTGGTGGTACCTTGTGTCACT 860  
 DB |||||  
 DB 867 GGTGACATCCAGATGCGACAGCTCGCAGAGGCGTGGCGGTGATCCGCTGTGAGAT 926  
 QY 861 GATCTCTTCTCGTGGCTTGGCATGTGCTGGTCTTGTGGCTGGCTGATCTTCTGCG 920  
 DB |||||  
 DB 927 GATGCTATGCTTACCGGTGCTCATCATCGGCTTGTGCTGGCGCATATCTTCCACT 986  
 QY 921 GGGAGCAAGTACCTGCGGGGTGTTTAACAGAGTCTGGGGATTGCCACAGCTCAT 980  
 DB |||||  
 DB 987 GGGTTTGTATCTGACCCCATCTCTGAGTCCCGGCTGACATCCAGACATATGCG 1046  
 QY 981 CATGGGCTACCACTTCACTTGTGCTGCTGCTTGGAGAGTCACTTCAATTTGTG 1035  
 DB |||||  
 DB 1047 CATTAACATGTGATGATGACATCTCTGCGATCATATGAGGGGATGCTGGGTGCTG 1101

RESULT 4  
 AAH25763  
 ID AAH25763 standard; cDNA; 1949 BP.  
 XX

AC AAH25763;

XX 14-AUG-2001 (first entry)

XX Oesophagus cancer associated DR2 coding sequence.

XX DR2; oesophagus cancer; ss.

XX Unidentified.

XX Key Location/Qualifiers

XX CDS 41..1480

XX FT /\*tag= a

XX FT /product= "DR2"

XX CN1283694-A.

XX PD 14-FEB-2001.

XX PF 10-AUG-1999; 99CN-0117523.

XX PR 10-AUG-1999; 99CN-0117523.

XX PA (ONCO-) ONCOLOGY INST TUMOR HOSPITAL CHINESE ACA.

XX PI Wang M, Xu Z, Xu X;

XX WP1; 2001-291757/31.

XX DR P-PSDB; AAB97000.

XX PT Gene associated with esophagus cancer -

XX PS Claim 4; Fig 1; 33pp; Chinese.

XX The present invention provides the protein and coding sequences of DR2,  
 CC which is associated with oesophagus cancer. Also provided is a method for  
 CC diagnosing diseases associated with an abnormal version of the nucleic  
 CC acid and encoded protein, involving detecting any nucleic acid sequence  
 CC mutations, methylation and variation at the RNA and polypeptide level.  
 CC The sequences can be used in the treatment of cancer.

XX  
 SQ Sequence 1949 BP; 382 A; 593 C; 530 G; 444 T; 0 other;

Query Match 13.2%; Score 165.4; DB 22; Length 1949;  
 Best Local Similarity 49.1%; Pred. No. 2.5e-37;  
 Matches 439; Conservative 0; Mismatches 456; Indels 0; Gaps 0;

```

QY 141 CTATCAAGTTGGCCCAAGATCTGACCGTGATGGCGCCATTGGCTTGGCTTCTCACTC 200
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 217 CTACCAAGCTTCCAGACGCGACGTGATGCTTCTGTGGGCTTCGGCTTCTCAAGAC 276
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 201 GAGTTCCGAGACACAGCTGAGCAAGTGGCCCTTCAACCTTTCATGCTGGCGCTTGG 260
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 277 TTTCCTGACAGGCTACGCGCTTCAAGCGCGTGGGCTTCAACTTCTGTGGAGGCTTCGG 336
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 261 TGTGCAATGGGCATCTCTGCTGACGGCTTCTCAAGCAGTCCCTTCTGGAGAGTGGT 320
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 337 CATCAAGTGGCGCTGCTCATGACAGGCTGGTCCACTTCTTACAGACCGCTCATCTGT 396
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 321 CATCAAGTGTTCAGTATTCGGCTGGCCACCATGATGCTTGTGCGTGCATCTCAGT 380
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 397 CGTGGGCTGGAAGACCTCATCAACGCTGACTTCTGCGTGGCCCTTGTCTGCGTGGCTT 456
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 381 GATGCTGCTTGGGGAAGTCAACTGGCGCAGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 440
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 457 TGGGCGAGTCTGGGTAAAGTCAAGCCCATTCAGCTGCTCATGACTTCTTCCAAGT 516
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 441 GACAGCTTTAGGCACTGAGGATGTCATGATATATCTTCAACACAGCTACCAT 500
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 517 GACCTCTTGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 576
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 501 GAACATGATGACATCTACGCTTTCGACGCTATTTTGGGCTGCTGCTGCTGCTGCTGCT 560
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 577 AGGCTCATACCATTCACACATTTGGCGCTTCTTGGGCTTCAAGTACCGGATCT 636
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 561 GCCAAGCTCTTACCCGAGGGAAGGATAAAGATCAACAGCAAGATACCATTT 620
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 637 CTACCAAGCACTAGAGAGAGAGAGAGAGAGAGAGAGAGATCTGTATACAGTGGACCT 696
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 621 GTCTCCATGCTGGCGCCCTTCTTGTGGATGTTCTGGCCAAAGTTTCAACTCTGCTCT 680
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 697 CTTTCATGATGATGACCCCTTCTGTGTGATGATGATGATGATGATGATGATGATGATG 756
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 681 GCTGAGAAGTCCAAATCGAAGAGAAATGCGGTGTTCAACACTATATGCTGTAGAGT 740
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 757 ATCTACCATGAGGAGACAGCCAGCCAGCCCATCAACACTTACTGCTCTTGGAGAC 816
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 741 CAGGTGTGTGACAGCCATCTCAGGATCATCTTGGCTCACCCCAAGGAGATCAGCA 800
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 817 CTGGGCTTACTTCTGGGCAATATCCAGTCCCTGACAAAGAGGCAAGCTGAGCAT 876
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 801 GACTTAATGTGACAGTGGGCTTGGCAGAGAGGCTGCTGTGGTACTTGTCTACCT 860
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 877 GGTGCAATCCAGATGCAAGCTGCGAGAGAGGAGTGGCCCTGGGTATCCGCTGTGAGAT 936
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 861 GATCCCTTCTCGTGGCTGTCATGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 920
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 937 GATCTCATGCTTACGATGCTTCAATCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 996
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 921 GGAAGCAAGTACCTGCGGGGCTTGTATACCGATGCTGGGATTTCCCAAGCTCAT 980
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 997 GGGTTTGTATACCTGAGCCCATCTCTGAGTCCCGGCTGACATCCAGACATGAGG 1056
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 981 CATGGGCTCAACTTCAAGCTTGTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1035
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1057 CATTAACAATCTGATGAGCATTCCTGGATCATAGGCGGCAATGCTGGTGTGTG 1111
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
  
```

RESULT 5  
 ABK49215  
 ID ABK49215 standard; cDNA; 1952 BP.  
 XX  
 AC ABK49215;  
 XX

DT 15-JUL-2002 (first entry)  
 XX

DE cDNA encoding human Rh type C gene (RHCG) protein.

KW RHCG: human; non-erythroid Rh type C glycoprotein;  
 XX chromosome 15q25; gene; ss.

OS Homo sapiens.

Key Location/Qualifiers  
 FH CDS 24..140  
 FT /\*tag= a  
 FT /product= "RHCG protein"  
 FT polyA\_signal 1885..1890  
 FT /\*tag= b

PN WO200220719-A2.

PD 14-MAR-2002.

PF 05-SEP-2001; 2001MO-US27503.

PR 07-SEP-2000; 2000US-230660P.

PA (NYBL-) NEW YORK BLOOD CENT INC.

PI Huang C, Liu Z;

DR WPI; 2002-351774/38.

DR P-PSDB; AAU78997.

PT Nucleic acid sequences encoding novel mammalian nonerythroid Rh type C  
 PT and glycoproteins which have a characteristic twelve transmembrane  
 PT domain structure -

PS Claim 2; Fig 1; 53pp; English.

XX This invention relates to the nucleic acid and protein sequences of  
 CC novel human and mouse non-erythroid Rh type C glycoprotein (RHCG).  
 CC The RHCG protein and the mouse homologue (rhcg) have a characteristic  
 CC 12 transmembrane domain structure and are expressed in kidneys and  
 CC testis. The invention also comprises a method for antibody that  
 CC specifically binds an epitope of the glycoprotein and a method for  
 CC detecting the protein using this antibody. The antibodies of the  
 CC invention may be used in Western blots, enzyme linked immunosorbent  
 CC assays (ELISA) or immunohistochemical assays to identify the non-  
 CC erythroid tissues, particularly kidney and testis, that express the  
 CC RHCG or Rhcg glycoproteins. The methods are used for detecting an  
 CC Rhcg or and Rhcg glycoprotein in a sample. The present sequence  
 CC represents the cDNA encoding the human RHCG Rh type C glycoprotein  
 CC (RHCG) protein sequence of the invention. The gene encoding this  
 CC protein is located on human chromosome 15q25.

XX  
 SQ Sequence 1952 BP; 398 A; 586 C; 524 G; 444 T; 0 other;

Query Match 13.2%; Score 165.4; DB 24; Length 1952;  
 Best Local Similarity 49.1%; Pred. No. 2.5e-37;  
 Matches 439; Conservative 0; Mismatches 456; Indels 0; Gaps 0;

```

QY 141 CTATCAAGTTGGCCCAAGATCTGACCGTGATGGCGCCATTGGCTTGGCTTCTCACTC 200
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 201 CTACCAAGCTTCCAGACGCGACGTGATGCTTCTGTGGGCTTCGGCTTCTCAAGAC 260
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 201 GAGTTCCGAGACACAGCTGAGCAAGTGGCCCTTCAACCTTTCATGCTGGCGCTTGG 260
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
  
```

Db 261 TTTCCTGAGGCGCTACGGCTTCAGGCGCGTGGCTTCACTTCTGTGGACAGCCTTCGG 320  
 Qy 261 TGTGCAAGTGGGCAATCCCTGCTGAGCGGCTTCCTGAGCCAGTTCCTTCCTGGAAAGTGGT 320  
 Db 321 CATCCAGTGGGCGCTGCTCATGCAAGGCTGGTTCACCTTTCACAAAGACCCCTACATGCT 380  
 Qy 321 CATCACTGTTCAGTATTCGGCTGGCCACCATGAGTCTTTCGTGGTGTGATCTCAGT 380  
 Db 381 CGTGGGCGTGGAGAACTCATCAACGCTGACTTTCGTGGCGCTGTGCTGGTGGCCTT 440  
 Qy 381 GGATGCTGCTTGGGGAAGTCACTTGGCCGATTTGGTGGTGGTGGTGGTGGAGT 440  
 Db 441 TGGGCACTTCCTGGGTAAAGTACGCCCATTCAGCTGCTCATCATGACTTTCCTCAAGT 500  
 Qy 441 GACAGCTTTAGGCAACTGAGAGTGTATCATGATATTCCTTCAACACAGACTACACAT 500  
 Db 501 GACCCCTTCCTGGTGAATGAGTTATTCCTTAACTGCTAAAGGTAAGAGATGACAG 560  
 Qy 501 GAACATGATGACATCTACGTGTTGCGAGCCTATTTTGGGCTGTCTGTGGCCTGGTCT 560  
 Db 561 AGGCTCCATGACATCCACACTTTGGCGCTACTTGGGCTCACAGTACCCGGATCCT 620  
 Qy 561 GCCAAAGCTCTACCCGAGGAGACGAGATTAAGATCAACAGACAGATACCATTT 620  
 Db 621 CTACCGAGCAACTGAGCAGAGAGAGAGAGACAGAAATTCGTGTAACAAGTGGACCT 680  
 Qy 621 GTTCGCAATGCTGGGCGCTCTTCTTGTGATGTTCTGGCCAACTTCAACTCTGCTCT 680  
 Db 681 CTTTCCCATGATGGCACCTCTTCTGTGATGATGATGATGATGATGATGATGATGAT 740  
 Qy 681 GCTGGAAGTCCATTCGAAAGAGAGATGCGTGTTCACACTTACATGCTGATGACAT 740  
 Db 741 ATCCATACCATGGGAGACACGACGACGAGCCGACATCAACACTGCTGCTTGGCAGC 800  
 Qy 741 CAGGTGTGACAGCATCTCAGGCTCATCTGCTGCTACACCCCAAGGAGATGAGAA 800  
 Db 801 CTGCTGCTTACCTCGTGGCAATTCAGTCCCTGCAAGAAAGGCAAGCTGAGCAT 860  
 Qy 801 GACTTATGTGACAGTGGGCTGTTGGAGAGAGCGCTGCTGTTGCTGCTGCTGCTGCT 860  
 Db 861 GGTGCAKATCCAGATGCGCAGCGCTGCGAGAGGGGTGGCCGTGGTACCGCTGAGAT 920  
 Qy 861 GATCCCTTCTCGTGGCTTGGCCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 920  
 Db 921 GATGCTCATGCTTACGCTGCGCTCATCATCGGCTTCGCTGCGGCAATCATCTCACCT 980  
 Qy 921 GGGAGCCAAAGTACCTGCGCGGGTGTGTAACCGAATGCTGGGAGATTCCCAAGCTCAT 980  
 Db 981 GGGTTTGTATACCTGACCCCATTCCTGAGATCCCGGCTGACATCCAGAGACATGTTG 1040  
 Qy 981 CATGGGCTTAACTTCACTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1035  
 Db 1041 CATTAACAATCGCATGCGATTCCTGGCATCATAGCGCGCATCGTGGTGTG 1095

RESULT 6  
 AACT6344  
 ID AACT6344 standard; cDNA; 1970 BP.

XX AACT6344;  
 XX 08-FEB-2001 (first entry)  
 DT Human ORF1899 polynucleotide sequence SEQ ID NO:3797.  
 XX  
 XX Human; open reading frame; ORF; detection; cytosolic; hepatotropic;  
 KM vulnervat; antiparkinsonian; neurotrophic; neuroprotective;  
 KM anticonvulsant; osteoplastic; antidiabetic; immunosuppressive; cardiatic;  
 KM immunosuppressant; thrombolytic; coagulant; vasodilator; antidiabetic;  
 KM hypotensive; dermatological; immunosuppressive; antidiabetic;  
 KM antiviral; antibacterial; antifungal; antineoplastic; antithyroid;  
 KM antidiabetic; gene therapy; cancer; proliferative disorder; hypertension;  
 KM neurodegenerative disorder; osteoarthritis; graft vs host disease;

KM cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;  
 KM cholesterol ester storage; systemic lupus erythematosus; infection;  
 KM severe combined immunodeficiency; malaria; autoimmune disorder; asthma;  
 KM allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;  
 KM bone damage; cartilage damage; antiinflammatory disease; coagulation;  
 KM thrombosis; contraceptive; ss.  
 OS Homo sapiens.  
 XX  
 XX WO200058473-A2.  
 PN  
 PD 05-OCT-2000.  
 XX  
 PF 31-MAR-2000; 2000WO-US08621.  
 XX  
 PR 31-MAR-1999; 99US-0127607.  
 PR 02-APR-1999; 99US-0127636.  
 PR 05-APR-1999; 99US-0127728.  
 PR 30-MAR-2000; 2000US-0540763.  
 XX  
 XX (CURA-) CURAGEN CORP.  
 XX  
 PI Shimketa RA, Leach M;  
 XX  
 DR WPI; 2000-602362/57.  
 DR P-PSDB; AAB42135.  
 XX  
 PT Novel nucleic acids and peptides derived from open reading frame X,  
 PT useful for treating e.g. cancers, proliferative disorders,  
 PT neurodegenerative disorders and cardiovascular disease -  
 XX  
 PS Claim 5; Page 2946-2948; 5507dp; English.  
 XX  
 CC AACT7446 to AACT7606 encode the proteins given in AAB40237 to AAB43397,  
 CC which represent the human ORF open reading frames 1 to 3161. The ORF  
 CC sequences have activities such as: cytosolic; hepatotropic; vulnervat;  
 CC antiparkinsonian; neurotrophic; neuroprotective;  
 CC osteoplastic; anticonvulsant; antidiabetic; immunosuppressive;  
 CC immunosuppressant; cardiatic; thrombolytic; coagulant; vasodilator;  
 CC antidiabetic; hypotensive; dermatological; immunosuppressive;  
 CC antiinflammatory; antibacterial; antiviral; antifungal; antineoplastic;  
 CC antithyroid; and antidiabetic. The sequences can be used for determining  
 CC the presence of or predisposition to, or preventing or treating  
 CC pathological conditions associated with an ORF-associated disorder. The  
 CC nucleic acids can be used to express ORF proteins in gene therapy  
 CC vectors. The proteins and nucleic acids may be used to treat cancers,  
 CC proliferative disorders, neurodegenerative disorders, osteoarthritis,  
 CC graft vs host disease, cardiovascular disease, diabetes mellitus,  
 CC hypertension, hypothyroidism, cholesterol ester storage, systemic lupus  
 CC erythematosus, severe combined immunodeficiency (SCID), AIDS, viral,  
 CC bacterial or fungal infection, malaria, autoimmune disorders, asthma,  
 CC allergies, aplastic anaemia, burns, wounds, bone and cartilage damage,  
 CC nocturnal haemoglobinuria, antiinflammatory disease; to enhance  
 CC coagulation; to inhibit thrombosis; and as a contraceptive.  
 CC  
 XX  
 SQ Sequence 1970 BP; 400 A; 593 C; 530 G; 445 T; 2 other;

Query Match 13.2%; Score 165.4; DB 21; Length 1970;  
 Best Local Similarity 49.1%; Pred. No. 2.5e-37;  
 Matches 439; Conservative 0; Mismatches 456; Indels 0; Gaps 0;

Qy 141 CTATCAAGTTGGCCAGATCTGACCGTATGCGCCATTTGGGCTTCTACCTC 200  
 Db 219 CTACCAAGCTTCAGAGACGTCAGCGATGCTTGTGGGCTTCGCTTCATGAC 278  
 Qy 201 GAGTTTCCGAGACACAGCTGAGCAGTGGCTTCACTTCAATGCTGGCGCTTG 260  
 Db 279 TTTCCTGAGGCGTACCGCTTCAAGCGCGTGGGCTTCAACTTCTTGGGCGCTTGG 338  
 Qy 261 TGTGCAAGTGGGCAATCCCTGCTGAGCGGCTTCCAGGACAGTTCCTTGGAGAGTGGT 320  
 Db 339 CATCCAGTGGGCGCTGCTCATGACGAGGCTGCTTCACTTTCACAAAGCCCTACATGCT 398





```

Db      629 CTACAGATCCAGCTGGAGAGACAGCATGCGCAGAGCTCTGTCTACAACTCTGACCT 688
Qy      621 GTCGGCATGCTGGGCGCCCTCTTGTGTGATGTTCTGGCGAAGTTCACTGCTCT 680
Db      689 CTTGGCCATGATCGGAGCATCTTCTGTGGGTTTCTGGCCGAGCTTCACTCCGCGCC 748
Qy      681 GCTGAGAAGTCAATCGAAGAAAGAAATGCGGTGTTCAACACTACTATGCTGTAGCAGT 740
Db      749 GACAGCGCTGGGGGATGGGGCAGCATCGGACCGTGGTCAACATATCTATTCACCTACCGC 808
Qy      741 CAGGTGTGTGACAGCCATCTCAGGCTCATCTTGGCTCACCCCCCAAGGAAAGATCAGCA 800
Db      809 AAGCACCCTCAGTACTTTCGCTTGTCTGAGCCCTTGTCTCAGTGAAGATGAGCATGGAACAT 868
Qy      801 GACTTATGTGACAGTGGGTGGTGGGAGAGGGGTGGCTGTGCTGCTGCTGCTGCTGCTGCT 860
Db      869 GGTCCACCTCCAGAACGAGACATGGCTGGAGAGTGTGTGGTGGGACATCAAGTGAACAT 928
Qy      861 GATCCCTTCTCTGCTGGCTGTCATGTCGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 920
Db      929 GATGCTGACACCTTTGGGGGCGCTGGCAGCTGGCTTCTGGCTGGGACTGTCTCCACACT 988
Qy      921 GGAAGCCAAATGACCTGCGGGGTGTGTAACTGAGTGTGGGATGCCCAAGCTCCAT 980
Db      989 GGGGTATTAAGTCTTTACGCTTATCTTGAATCCAGATTAACTGCAAGACACATGTGG 1048
Qy      981 CATGGGCTACAACTTCACTGCTGGGTCTGCTGGAGAGATCATCTGCTGCTGCT 1040
Db      1049 TGTTCACAACTTCCATGGGAATGACAGGGGCTCTGGGGCCATCTGGAGTGTGATGGC 1108
Qy      1041 GGTCCTTGATACCGTGGAGC 1061
Db      1109 TGCACCTGGCCACCCAGCAAGC 1129

RESULT 8
AAH18728/c
ID      AAH18728 standard; cDNA; 3706 BP.
XX
AC      AAH18728;
DT      26-JUN-2001 (first entry)
XX
DE      Human cDNA sequence SEQ ID NO:19002.
XX
KM      Human; primer: detection; diagnosis; antisense therapy; gene therapy; ss.
XX
OS      Homo sapiens.
XX
PN      EP1074617-A2.
XX
PD      07-FEB-2001.
XX
PF      28-JUL-2000; 2000EP-0116126.
XX
PR      29-JUL-1999; 99JP-0248036.
PR      27-AUG-1999; 99JP-0300253.
PR      11-JAN-2000; 2000JP-0118776.
PR      02-MAY-2000; 2000JP-0183767.
PR      09-JUN-2000; 2000JP-0241899.
XX
PA      (HELI-) HELIX RES INST.
XX
PI      Oca T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI      Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX
DR      WPI; 2001-318749/34.
XX
PT      Primer sets for synthesizing polynucleotides, particularly the 5602
PT      full-length cDNAs defined in the specification, and for the detection
PT      and/or diagnosis of the abnormality of the proteins encoded by the
PT      full-length cDNAs.
XX

```

```

PS      Claim 8; SEQ ID 19002; 2537pp + CD ROM; English.
XX
CC      The present invention describes primer sets for synthesizing 5602
CC      full-length cDNAs defined in the specification. Where a primer set
CC      comprises: (a) an oligo-dT primer and an oligonucleotide complementary
CC      to the complementary strand of a polynucleotide which comprises one of
CC      the 5602 nucleotide sequences defined in the specification, where the
CC      oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC      of an oligonucleotide comprising a sequence complementary to the
CC      complementary strand of a polynucleotide which comprises a 5'-end
CC      sequence and an oligonucleotide comprising a sequence complementary to a
CC      polynucleotide which comprises a 3'-end sequence, where the
CC      oligonucleotide comprises at least 15 nucleotides and the combination of
CC      the 5'-end sequence/3'-end sequence is selected from those defined in
CC      the specification. The primer sets can be used in antisense therapy and
CC      in gene therapy. The primers are useful for synthesizing polynucleotides,
CC      particularly full-length cDNAs. The primers are also useful for the
CC      detection and/or diagnosis of the abnormality of the proteins encoded by
CC      the full-length cDNAs. The primers allow obtaining of the full-length
CC      cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC      AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to
CC      AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632
CC      represent oligonucleotides, all of which are used in the exemplification
CC      of the present invention.
XX
SQ      Sequence 3706 BP; 883 A; 1050 C; 684 G; 1089 T; 0 other;
Query Match      12.2%; Score 152.4; DB 22; Length 3706;
Best Local Similarity 99.4%; Pred. No. 2.1e-33;
Matches 153; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy      333 CAGATTGCGGTGGCCACCATGATGCTTGTCCGTGCTGATCTCAGTGAATGCTGCTT 392
Db      2906 CAGTATTCGGCTGGCCACCATGATGCTTGTCCGTGCTGATCTCAGTGAATGCTGCTT 2847
Qy      393 GGGGAAGTCACTTGGCGCAGTTGGTGTGATGCTGCTGCTGAGAGTGAACGCTTTAGG 452
Db      2846 GGGGAAGTCACTTGGCGCAGTTGGTGTGATGCTGCTGCTGAGAGTGAACGCTTTAGG 2787
Qy      453 CAACCTGAGATGTCATCACTAATATCTTCAAC 486
Db      2786 CAACCTGAGATGTCATCACTAATATCTTCAAC 2753

RESULT 9
ABK49216
ID      ABK49216 standard; cDNA; 2098 BP.
XX
AC      ABK49216;
DT      15-JUL-2002 (first entry)
XX
DE      cDNA encoding mouse Rh type C gene (rhcg) protein.
XX
KM      RHCG: mouse; non-erythroid Rh type C glycoprotein;
KM      chromosome 7; gene; ss.
XX
OS      Mus sp.
XX
FH      Key
FH      CDS      Location/Qualifiers
FT      FT      /*tag= a
FT      FT      /product= "RhCG protein"
FT      FT      1934..1939
FT      FT      /*tag= b
XX
PN      WO200220719-A2.
XX
PD      14-MAR-2002.
XX
PF      05-SEP-2001; 2001WO-US27503.
XX
PR      07-SEP-2000; 2000US-230660P.
XX

```

XX (NYBL-) NEW YORK BLOOD CENT INC.  
 XX  
 XX Huang C, Liu Z;  
 XX WPI: 2002-351774/38.  
 DR P-PSDB; AAU78998.  
 XX  
 PT Nucleic acid sequences encoding novel mammalian nonerythroid Rh type C  
 PT and glycoproteins which have a characteristic twelve transmembrane  
 PT domain structure -  
 PS  
 PS Claim 3; Fig 1; 53pp; English.  
 XX  
 XX This invention relates to the nucleic acid and protein sequences of  
 CC novel human and mouse non-erythroid Rh type C glycoprotein (RhCG).  
 CC The RhCG protein and the mouse homologue (rhcg) have a characteristic  
 CC 12 transmembrane domain structure and are expressed in kidneys and  
 CC testis. The invention also comprises a method for antibody that  
 CC specifically binds an epitope of the glycoprotein and a method for  
 CC detecting the protein using this antibody. The antibodies of the  
 CC invention may be used in Western blots, enzyme linked immunosorbent  
 CC assays (ELISA) or immunohistochemical assays to identify the non-  
 CC erythroid tissues, particularly kidney and testis, that express the  
 CC RhCG or Rhcg glycoproteins. The methods are used for detecting an  
 CC Rhcg or and RhCG glycoprotein in a sample. The present sequence  
 CC represents the cDNA encoding the mouse RhCG Rh type C glycoprotein  
 CC (RhCG) protein sequence of the invention. The gene encoding this  
 CC protein is located on murine chromosome 7.  
 XX  
 XX Sequence 2098 BP; 458 A; 594 C; 559 G; 487 T; 0 other;  
 SQ  
 Query Match 12.0%; Score 150.2; DB 24; Length 2098;  
 Best Local Similarity 47.5%; Pred. No. 6.8e-33;  
 Matches 446; Conservative 0; Mismatches 493; Indels 0; Gaps 0;

DB 798 ACCCTCTTCTTGAGATATACGTGCGCCAGTTCAATTCACCAAGTTCTTCCACGGAGAT 857  
 QY 697 GAAAGAAAGATGCCGTTCAACACCTACTATGCTGTAGACGACGTTGTACAGCC 756  
 DB 858 GCCCAGACCGAGACGCCCTCAATACCTACTCTCTTGGCAGGATGTCTAACCA 917  
 QY 757 ATCTCAGGGTCAATCTTGGCTCAACCCCAAGGAAAGATCAGACAGATTATGTGCA 816  
 DB 918 GTACAGATATCCAGTATTTGTAACAAAGAGGCAAGTTGATATGTGACATCCAGAT 977  
 QY 817 GCGGTGTTGCGAGAGCGCTGTGGTGTGCTGTGTAACCTGATACCTTCTTCGTGG 876  
 DB 978 GCCACGCTTCAGAGTGGGTGGGTGGGACAGCTCGAGATGATGCTTCAACACTTAC 1037  
 QY 877 CTTCGATGCGTGGGTCTTGGGCTGATCTCCGTCGGGAGAGCCAAATCTCG 936  
 DB 1038 GGCCTCTCATCTGGGTCTTCTGTGGCATTTTCTCCACCTTAGAATTTGCATACCTA 1097  
 QY 937 CCGGGGTGTTGTAACCAAGTCTGGGATTTCCACACGCTTCATCAGGCTTACAACTTC 996  
 DB 1098 ACGCATTCCTGAGAGTCCCGCTTCGATCCAGACACATGTGCTTCACTTCAACCTGCAC 1157  
 QY 997 AGCTTGCTGGTCTCTTGGAGAGATCATCTTACATTGTG 1035  
 DB 1158 GGCATTCCTGCAATCATAGCGCGCATTTGGTGTGTG 1196  
 RESULT 10  
 ABA42886/c  
 ID ABA42886 standard; DNA; 486 BP.  
 XX  
 XX ABA42886;  
 AC  
 AC 01-FEB-2002 (first entry)  
 DT  
 XX Human breast cell single exon nucleic acid probe #1581.  
 XX Human; microarray; single exon probe; gene expression; breast;  
 KW disease; cancer; ss.  
 OS Homo sapiens.  
 XX  
 XX MO200157271-A2.  
 PN  
 PD 09-AUG-2001.  
 XX  
 XX 30-JAN-2001; 2001MO-US00662.  
 PF  
 XX 04-FEB-2000; 2000US-0180312.  
 PR 26-MAY-2000; 2000US-0207456.  
 PR 30-JUN-2000; 2000US-0608408.  
 PR 03-AUG-2000; 2000US-0632366.  
 PR 21-SEP-2000; 2000US-0234687.  
 PR 27-SEP-2000; 2000US-0236359.  
 PR 04-OCT-2000; 2000GB-0024263.  
 XX  
 XX (MOLE-) MOLECULAR DYNAMICS INC.  
 PA Penn SG, Hanzel DK, Chen W, Rank DR;  
 PI WPI; 2001-496933/54.  
 DR  
 XX  
 XX New spatially-addressable set of single exon nucleic acid probes,  
 PT useful for measuring gene expression in sample derived from human  
 PT breast, comprises number of single exon nucleic acid probes  
 PT  
 PS Claim 1; SEQ ID NO 1581; 327bp + sequence listing; English.  
 XX  
 XX The invention relates to a spatially-addressable set of single exon  
 CC nucleic acid probes for measuring gene expression in a sample derived  
 CC from human breast and BT 474 cells. The method involves contacting  
 CC the probes with a collection of detectably labelled nucleic acids





Query Match 11.1%; Score 139.8; DB 22; Length 486;  
Best Local Similarity 92.5%; Pred. No. 3.2e-30;  
Matches 147; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY 477 TATCTTCAACAGACACTACCAATGAACATGATGACATCTACGTGTTCGAGCCTATT 536  
DB 458 TTTATTGACAGACACTACCAATGAACATGATGACATCTACGTGTTCGAGCCTATT 399  
QY 537 TGGGCTGTCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGG 596  
DB 398 TGGGCTGTCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGG 339

QY 597 TCAGACAGCAACGATATACCCAGTTTGTCTGCCATGCTGGG 635  
DB 338 TCAGAGAGCAACGATATACCCAGTTTGTCTGCCATGCTGGG 300

RESULT 15  
AA11615/c  
ID AA11615 standard; DNA; 486 BP.  
XX  
AC AA11615;  
XX  
DT 12-OCT-2001 (first entry)  
XX  
DE Probe #1548 for gene expression analysis in human cervical cell sample.  
XX  
KM Probe; human; microarray; gene expression; cervical epithelial cell;  
XX  
KW cervical cancer; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO200157278-A2.  
XX  
PD 09-AUG-2001.  
XX  
PF 30-JAN-2001; 2001WO-US00670.  
XX  
PR 04-FEB-2000; 2000US-0180312.  
PR 26-MAY-2000; 2000US-0207456.  
PR 30-JUN-2000; 2000US-0608408.  
PR 03-AUG-2000; 2000US-0632366.  
PR 21-SEP-2000; 2000US-0234687.  
PR 27-SEP-2000; 2000US-0236359.  
PR 04-OCT-2000; 2000GB-0024263.  
XX  
PA (MOLE-) MOLECULAR DYNAMICS INC.  
XX  
PI Penn SG, Hanzel DK, Chen W, Rank DR;  
XX  
DR WPI; 2001-488901/53.  
XX  
PT Human genome-derived single exon nucleic acid probes useful for  
XX  
PS analyzing gene expression in human cervical epithelial cells -  
XX  
SQ Claim 25; SEQ ID No 1548; 487bp; English.

CC The present invention relates to human single exon nucleic acid probes  
CC (SENPs). The present sequence is one such probe. The SENPs are derived  
CC from human HeLa cells. The SENPs can be used to produce a single exon  
CC microarray, which can be used for measuring human gene expression in a  
CC sample derived from human cervical epithelial cells. By measuring gene  
CC expression, the probes are therefore useful in grading and/or staging  
CC of diseases of the cervix, notably cervical cancer.  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.

SQ Sequence 486 BP; 137 A; 112 C; 112 G; 125 T; 0 other;

Query Match 11.1%; Score 139.8; DB 22; Length 486;  
Best Local Similarity 92.5%; Pred. No. 3.2e-30;  
Matches 147; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY 477 TATCTTCAACAGACACTACCAATGAACATGATGACATCTACGTGTTCGAGCCTATT 536  
DB 458 TTTATTGACAGACACTACCAATGAACATGATGACATCTACGTGTTCGAGCCTATT 399  
QY 537 TGGGCTGTCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGG 596  
DB 398 TGGGCTGTCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGG 339

QY 597 TCAGACAGCAACGATATACCCAGTTTGTCTGCCATGCTGGG 635  
DB 338 TCAGAGAGCAACGATATACCCAGTTTGTCTGCCATGCTGGG 300

Search completed: April 8, 2003, 04:00:17  
Job time : 355 secs

GenCore version 5.1.4 p5\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 7, 2003, 19:53:30 ; Search time 2036 Seconds  
(without alignments)  
9975.020 Million cell updates/sec

Title: US-09-600-714-41

Perfect score: 1254  
Sequence: 1 atgagctcaagaccgcg.....attgcctgtgatttaa 1254

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database :

EST:\*  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estlin:\*  
4: em\_estmu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_hic:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_hic:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: gb\_gss:\*  
18: em\_gss\_hum:\*  
19: em\_gss\_inv:\*  
20: em\_gss\_pln:\*  
21: em\_gss\_vrt:\*  
22: em\_gss\_fun:\*  
23: em\_gss\_mam:\*  
24: em\_gss\_mus:\*  
25: em\_gss\_other:\*  
26: em\_gss\_pro:\*  
27: em\_gss\_rtd:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	398.6	31.8	683	14	BM719724 UI-E-EJ1-AL699045
2	392.4	31.3	443	9	AL699045 DKFZP686M
3	343.2	27.4	551	13	BM254191 515553 MA
4	317.6	25.3	643	14	BM684087 UI-E-EJ1-N59044
5	302.2	24.1	434	14	N59044 yv60b12.r1
6	294.2	23.5	499	12	BF603905 269531 MA

7	288.2	23.0	468	12	BF602079	BF602079 267224 MA
8	271.2	21.6	468 <td>13 <td>BG944752</td> <td>BG944752 ax54409.x</td> </td>	13 <td>BG944752</td> <td>BG944752 ax54409.x</td>	BG944752	BG944752 ax54409.x
9	270.4	21.6	468 <td>14 <td>R10548</td> <td>R10548 yf3e01.r1</td> </td>	14 <td>R10548</td> <td>R10548 yf3e01.r1</td>	R10548	R10548 yf3e01.r1
10	253.6	20.2	343 <td>14 <td>T84327</td> <td>T84327 yd37a05.r1</td> </td>	14 <td>T84327</td> <td>T84327 yd37a05.r1</td>	T84327	T84327 yd37a05.r1
11	243	19.4	515 <td>12 <td>BF191605</td> <td>BF191605 239237 MA</td> </td>	12 <td>BF191605</td> <td>BF191605 239237 MA</td>	BF191605	BF191605 239237 MA
12	236.8	18.9	461	9 <td>AA259801</td> <td>AA259801 v887911.x</td>	AA259801	AA259801 v887911.x
13	231.8	18.3	446 <td>10 <td>BB847808</td> <td>BB847808 BB847808</td> </td>	10 <td>BB847808</td> <td>BB847808 BB847808</td>	BB847808	BB847808 BB847808
14	229.8	18.5	454	12 <td>BG087038</td> <td>BG087038 H3134C09-AA01691 mg90e03.x</td>	BG087038	BG087038 H3134C09-AA01691 mg90e03.x
15	208.6	16.6	431	9 <td>AA016691</td> <td>AA016691 mg90e03.x</td>	AA016691	AA016691 mg90e03.x
16	208	16.6	431	10 <td>BB849234</td> <td>BB849234 BB849234</td>	BB849234	BB849234 BB849234
17	199.6	15.9	430 <td>10 <td>BB850227</td> <td>BB850227 BB850227</td> </td>	10 <td>BB850227</td> <td>BB850227 BB850227</td>	BB850227	BB850227 BB850227
18	196.2	15.6	350 <td>10 <td>BB842859</td> <td>BB842859 BB842859</td> </td>	10 <td>BB842859</td> <td>BB842859 BB842859</td>	BB842859	BB842859 BB842859
19	194.6	15.5	368 <td>14 <td>T89255</td> <td>T89255 yd37a05.s1</td> </td>	14 <td>T89255</td> <td>T89255 yd37a05.s1</td>	T89255	T89255 yd37a05.s1
20	192.8	15.4	445 <td>10 <td>BB849924</td> <td>BB849924 BB849924</td> </td>	10 <td>BB849924</td> <td>BB849924 BB849924</td>	BB849924	BB849924 BB849924
21	190.6	15.2	588 <td>13 <td>B1546440</td> <td>B1546440 603188741</td> </td>	13 <td>B1546440</td> <td>B1546440 603188741</td>	B1546440	B1546440 603188741
22	190	15.2	386 <td>10 <td>BB842875</td> <td>BB842875 BB842875</td> </td>	10 <td>BB842875</td> <td>BB842875 BB842875</td>	BB842875	BB842875 BB842875
23	188.6	15.0	382 <td>10 <td>BB842470</td> <td>BB842470 BB842470</td> </td>	10 <td>BB842470</td> <td>BB842470 BB842470</td>	BB842470	BB842470 BB842470
24	188	15.0	645 <td>10 <td>AW553137</td> <td>AW553137 L0222E01-AW008551 mg86a09.x</td> </td>	10 <td>AW553137</td> <td>AW553137 L0222E01-AW008551 mg86a09.x</td>	AW553137	AW553137 L0222E01-AW008551 mg86a09.x
25	186.2	14.8	370	9 <td>AA008551</td> <td>AA008551 mg86a09.x</td>	AA008551	AA008551 mg86a09.x
26	183.4	14.6	466 <td>12 <td>BF557252</td> <td>BF557252 UI-R-CO-G</td> </td>	12 <td>BF557252</td> <td>BF557252 UI-R-CO-G</td>	BF557252	BF557252 UI-R-CO-G
27	183.2	14.6	456 <td>12 <td>BB842399</td> <td>BB842399 BB842399</td> </td>	12 <td>BB842399</td> <td>BB842399 BB842399</td>	BB842399	BB842399 BB842399
28	179.4	14.3	378 <td>10 <td>BB842399</td> <td>BB842399 BB842399</td> </td>	10 <td>BB842399</td> <td>BB842399 BB842399</td>	BB842399	BB842399 BB842399
29	178.2	14.2	482	9 <td>AI632247</td> <td>AI632247 t220f10.x</td>	AI632247	AI632247 t220f10.x
30	176.8	14.1	335 <td>12 <td>BG221305</td> <td>BG221305 RST41110</td> </td>	12 <td>BG221305</td> <td>BG221305 RST41110</td>	BG221305	BG221305 RST41110
31	176.8	14.1	815 <td>12 <td>BG197922</td> <td>BG197922 RST17171</td> </td>	12 <td>BG197922</td> <td>BG197922 RST17171</td>	BG197922	BG197922 RST17171
32	176.2	14.1	791 <td>12 <td>BG181882</td> <td>BG181882 RST736 AC</td> </td>	12 <td>BG181882</td> <td>BG181882 RST736 AC</td>	BG181882	BG181882 RST736 AC
33	173.6	13.8	395 <td>12 <td>BG182395</td> <td>BG182395 RST1261 A</td> </td>	12 <td>BG182395</td> <td>BG182395 RST1261 A</td>	BG182395	BG182395 RST1261 A
34	173	13.8	365 <td>12 <td>BF510577</td> <td>BF510577 UI-H-B1A-BG219864 RST39631</td> </td>	12 <td>BF510577</td> <td>BF510577 UI-H-B1A-BG219864 RST39631</td>	BF510577	BF510577 UI-H-B1A-BG219864 RST39631
35	171.2	13.7	415 <td>12 <td>BG219864</td> <td>BG219864 RST39631</td> </td>	12 <td>BG219864</td> <td>BG219864 RST39631</td>	BG219864	BG219864 RST39631
36	170.2	13.6	357 <td>12 <td>BG191289</td> <td>BG191289 RST10377</td> </td>	12 <td>BG191289</td> <td>BG191289 RST10377</td>	BG191289	BG191289 RST10377
37	168.6	13.4	358 <td>10 <td>BB841796</td> <td>BB841796 BB841796</td> </td>	10 <td>BB841796</td> <td>BB841796 BB841796</td>	BB841796	BB841796 BB841796
38	167.6	13.4	492 <td>14 <td>N53959</td> <td>N53959 yv60b12.s1</td> </td>	14 <td>N53959</td> <td>N53959 yv60b12.s1</td>	N53959	N53959 yv60b12.s1
39	165.8	13.2	355 <td>12 <td>BG208730</td> <td>BG208730 RST28364</td> </td>	12 <td>BG208730</td> <td>BG208730 RST28364</td>	BG208730	BG208730 RST28364
40	165.8	13.2	768 <td>12 <td>BG186163</td> <td>BG186163 RST5008 A</td> </td>	12 <td>BG186163</td> <td>BG186163 RST5008 A</td>	BG186163	BG186163 RST5008 A
41	165.4	13.2	386 <td>10 <td>BB842372</td> <td>BB842372 BB842372</td> </td>	10 <td>BB842372</td> <td>BB842372 BB842372</td>	BB842372	BB842372 BB842372
42	165.2	13.2	372 <td>10 <td>BB844571</td> <td>BB844571 BB844571</td> </td>	10 <td>BB844571</td> <td>BB844571 BB844571</td>	BB844571	BB844571 BB844571
43	165.2	13.2	1864	11 <td>BC015461</td> <td>BC015461 Homo sapi-</td>	BC015461	BC015461 Homo sapi-
44	163.4	13.0	377	12 <td>BG186607</td> <td>BG186607 RST5576 A</td>	BG186607	BG186607 RST5576 A
45	162	12.9	476	12 <td>BG214121</td> <td>BG214121 RST33749</td>	BG214121	BG214121 RST33749

## ALIGNMENTS

RESULT 1  
LOCUS BM719724 683 bp mRNA linear EST 01-MAR-2002  
DEFINITION UI-E-EJ1-aj1-i-11-0-UI.r1 UI-E-EJ1 Homo sapiens cDNA clone  
ACCESSION UI-E-EJ1-aj1-i-11-0-UI 5', mRNA sequence.  
VERSION BM719724  
KEYWORDS BM719724.1 GI:19038503  
SOURCE EST.  
ORGANISM human.  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 683)  
AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.  
TITLE Normalization and subtraction: two approaches to facilitate gene  
JOURNAL discovery  
MEDLINE Genome Res. 6 (9), 791-806 (1996)  
COMMENT 97044477  
Contact: Soares, MB  
Program for Rat Gene Discovery and Mapping  
University of Iowa  
451 Eckstein Medical Research Building Iowa City, IA 52242, USA  
Tel: 319 335 8250  
Fax: 319 335 9565  
Email: msocares@blue.weeg.uiowa.edu  
Tissue Procurement: Dr. Gregg Hageman  
CDNA Library Preparation: Dr. M. Bento Soares, University of Iowa  
CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa  
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa





```

RESULT 3
LOCUS      BM254191                    551 bp    mRNA    linear    EST 17-DEC-2001
DEFINITION 515552 MARC 3BOV Bos taurus cDNA 5', mRNA sequence.
ACCESSION  BM254191
VERSION     BM254191.1    GI:17889790
KEYWORDS   EST.
SOURCE     Bos taurus
            cow.
ORGANISM   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidea;
            Bovidae; Bovinae; Bos.
REFERENCE  1 (bases 1 to 551)
            Smith,T.P.L., Grosse,W.M., Freking,B.A., Roberts,A.J., Stone,R.T.,
            Casas,E., Wray,J.E., White,J., Cho,J., Fahrrenkrug,S.C., Bennett
            ,G.L., Heaton,M.P., Laegreid,W.W., Rohrer,G.A., Chikko-Mckown,C.G.,
            Perta,G., Holt,I., Karanymcheva,S., Liang,F., Quackenbush,J. and
            Keele,J.W.
            Sequence evaluation of four pooled-tissue normalized bovine cDNA
            libraries and construction of a gene index for cattle
            Genome Res. 11 (4), 626-630 (2001)
TITLE      Contact: Smith TPL
JOURNAL    USDA, ARS, US Meat Animal Research Center
MEDLINE    PO Box 166, Clay Center, NE 68933-0166, USA
COMMENT    Tel: 402 762 4366
            Fax: 402 762 4390
            Email: smith@email.marc.usda.gov
            Single pass sequencing. Bases called and alt trimmed with phred
            v0.960904.e. Vector identified by cross_match with the -minscore 18
            and -mismatch 12 options.
            PCR Primers
            FORWARD: AGGAACAGCTATGACCAT
            BACKWARD: GTTTCGCCAGTCACGACG
            Plate: 113 row: C column: 17
            Seq primer: ATTTAGTGACACTATG.
            Location/Qualifiers
                source
                1. 551
                /organism="Bos taurus"
                /db_xref="taxon:9913"
                /clone_lib="MARC 3BOV"
                /tissue_type="pooled"
                /lab_host="DH10B"
                /note="Vector: pCMV SPORT6; Site 1: NotI; Site 2: SalI;
                library made from pooled tissue from marrow, alveolar
                macrophage, ovary, fetal semitendinosus muscle, and fetal
                longissimus muscle."
BASE COUNT 117 a 148 c 149 g 137 t
ORIGIN
Query Match 27.4%; Score 343.2; DB 13; Length 551;
Best Local Similarity 76.6%; Pred. No. 2.7e-86;
Matches 420; Conservative 0; Mismatches 128; Indels 0; Gaps 0;
QY 346 GCCACCATGAGTGGTCTTGTGCGTCTGATCTCACTGATGTCGTCTTGGGGAAGTCAAC 405
Db 2 GCCACCATGATCGCTAATCTGTGTATCTCCGGGGTGTGTCCTGGGGAAGTCAAC 61
QY 406 TTGGCGCATGTTGGTGTGATGTGTGTGTGAGAGTGACAGCTTTAGGCAACCTTAGAGATG 465
Db 62 CTGTGTGAGCGGTGATATATACCTGATAGAGGTGACAGCTTCAAGTCCACAGAGGCTG 121
QY 466 GTCAATCATATATCTTCAACACAGACTACACATGATGATGATGATGATGATGATGATGATG 525
Db 122 GTCAATCATATCTCTGATATGATGATGATGATGATGATGATGATGATGATGATGATGATG 181
QY 526 GCACCTATTTTGGGCTGTCTGTGGCTGTGCTGCTGCCAAAGCCTTACCCGAGGGAACG 585
Db 182 GCGGCTATTTTGGGCTGTGCTGTGCTGCTGCCAAAGCCTTACCCGAGGGAACG 241
QY 586 GAGATTAAGTTCAGACGACGACGATACCCAGTTTGTCTGCGATGCTGGGCGCCCTCTTC 645
Db 242 GAGACAAAGTTCAGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 301

```

```

QY 646 TTGTGATGTTTGGCCCAAGTTTCACTGTCTGTGAGAAATCCAAAGGAAG 705
Db 302 TTGTGATATTTGGCCGAGTTTCACTGTCTGTGAGAAATCCAAAGGAAG 361
QY 706 AATGCCGTTCACACCTTACTATGTGTAGACAGTACCGTGTGACAGCCATTCAGGG 765
Db 362 ATGCCCGTTCACACCTTACTATGTGTAGACAGTACCGTGTGACAGCCATTCATG 421
QY 766 TCATCTTGGCTACCCCAAGGAGATGACAGAACTTATGTGACAGTGGCGTGTG 825
Db 422 TCAGCTTGGCTACCCCAAGGAGATGATGTGACTACATCCACAAAGGACAGTCTG 481
QY 826 GCAGAGCGCTGGCTGTGGTGTACCTCGTGTACCTGATCCCTTCCGCTGGCTTGCATG 885
Db 482 GCAGAGAGTGTGGTGTGGTGTGCGGTCCTTCTTACTGATGATGATGATGATGATGATG 541
QY 886 GTGCTGGG 893
Db 542 GTGCTGGG 549

RESULT 4
LOCUS      BM684087/c                    643 bp    mRNA    linear    EST 27-FEB-2002
DEFINITION UI-E-EJ1-aj1-1-11-0-UI.s1 UI-E-EJ1 Homo sapiens cDNA clone
ACCESSION  BM684087
VERSION     BM684087.1    GI:18993983
KEYWORDS   EST.
SOURCE     Homo sapiens
            human.
ORGANISM   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE  1 (bases 1 to 643)
            Bonaldo,M.F., Lennon,G. and Soares,M.B.
            Normalization and subtraction: two approaches to facilitate gene
            discovery
            Genome Res. 6 (9), 791-806 (1996)
JOURNAL    97044477
MEDLINE    Contact: Soares, MB
            Program For Rat Gene Discovery and Mapping
            University of Iowa
            451 Eckstein Medical Research Building Iowa City, IA 52242, USA
            Tel: 319 335 8250
            Fax: 319 335 9565
            Email: msoares@blue.weeg.uiowa.edu
            Tissue Procurement: Dr. Gregg Hageman
            CDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
            DNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
            Clone Distribution: Researchers may obtain clones from Research
            Genetics (www.resgen.com).
            The following repetitive elements were found in this cDNA
            sequence: 1-26, >AT richlow_complexity
            Seq primer: M13 Forward
            POLYA=yes.
            Location/Qualifiers
                source
                1. 643
                /organism="Homo sapiens"
                /db_xref="taxon:9606"
                /clone_lib="UI-E-EJ1-aj1-1-11-0-UI"
                /clone_lib="UI-E-EJ1"
                /tissue_type="fetal eyes, lens, eye anterior segment,
                optic nerve, retina, Retina foveal and Macular, RPE and
                Choroid"
                /dev_stage="fetal and adult"
                /lab_host="DH10B (Life Technologies) (T1 phage resistant)"
                /note="Organ: eye; Vector: pRTT3-Pac (Pharmacia) with a
                modified polylinker; Site 1: EcoR I; Site 2: Not I;
                UI-E-EJ1 is a subcloned cDNA library constructed
                according to Bonaldo, Lennon and Soares, Genome Research,
                6:791-806, 1996. First strand cDNA synthesis was primed

```



TITLE  
JOURNAL  
MEDLINE  
COMMENT

Casas, E., Wray, J.E., White, J., Cho, J., Fahrenkrug, S.C., Bennett, G.L., Heaton, M.P., Laegreid, W.W., Rohrer, G.A., Chitko-McKown, C.G., Perte, G., Holt, I., Karamycheva, S., Liang, F., Quackenbush, J., and Keeler, J.W.

Sequence evaluation of four pooled-tissue normalized bovine cDNA libraries and construction of a gene index for cattle

Genome Res. 11 (4), 626-630 (2001)

21180013

Contact: Smith TPL  
USDA, ARS, US Meat Animal Research Center  
PO Box 166, Clay Center, NE 68933-0166, USA  
Tel: 402 762 4366  
Fax: 402 762 4390  
Email: smith@email.marc.usda.gov

Single pass sequencing. Bases called and alt trimmed with phred v0.980904.e. Vector identified by cross\_match with the -minscore 18 and -mismatch 12 options.

PCR Primers  
FORWARD: AGGAACAGCTATGACCAT  
BACKWARD: GTTTCCTCAGTCAGCAGC  
Plate: 50 row: D column: 2  
Seq primer: ATTAGGTGACACTATAG.

FEATURES  
source  
1. .499  
/organism="Bos taurus"  
/db\_xref="taxon:9913"  
/clone\_1lb="MARC 3BOV"  
/tissue\_type="pooled"  
/lab\_host="DH10B"  
/note="Vector: PCMV SPORT6; Site\_1: NotI; Site\_2: SalI; Library made from pooled tissue from marrow, alveolar macrophage, ovary, fetal semitendinosus muscle, and fetal longissimus muscle."

BASE COUNT 101 a 134 c 133 g 131 t

ORIGIN

Query Match 23.5%; Score 294.2; DB 12; Length 499;  
Best Local Similarity 74.9%; Pred. No. 2.2e-72;  
Matches 368; Conservative 0; Mismatches 123; Indels 0; Gaps 0;

194 TCACCTGAGCTTCCGAGACACAGCTGAGCAGTGTGGCTTCACTCTTCATGCTGG 253  
1 TCATATCTCTTTCGCGAGACATGCTGAGCAGATGCTTCAATCTTTCGCTGG 60  
254 CGCTTGATGTCAGTGGGAGATCCGCTGAGCGGCTTCTGAGCAAGTTCCTGCGGA 313  
61 TCTTCGGGGTCACTGAGCAGCCCTGCTGAGTGGCTTCTGACCACTTCCTTGAAGA 120  
314 AGGTGTCATCACATGTTCAATTCGCTGGCCACCATGATGCTTGTGCTGCTGA 373  
121 AGATGTCATGATATGTCAGTATTCAGAGAGCCACCATGATGCTATCTGTGCTGA 180  
374 TCTAGTGAATGCTGCTTGGGGAAGTCACTTGGCCAGTTGGTGTGATGCTGG 433  
181 TCTCCGGGGTGTGCTCTGGGGAAGTCACTGCTGACGTGTGATTAAGCGCTGA 240  
434 TGAAGTACAGCTTTAGGCAACCTGAGAGTGTATGATTAATCTTCAACAGACT 493  
241 TAGAGTACAGCTTTAGTGTGACGAGGCTGTGATTAAGTCACTTCAATGAACA 300  
494 ACCACATGATGATGACATCTAAGTGTGACGCTTATTTGGGCTGTGCTGGCT 553  
301 TGATGTAAGCATGATGATATGACATGATGAGGCTTATTTGGGCTGTGCTGGCT 360  
554 GGTGCTTCCAAAGCTTTACCCGAGGAGAGAGATTAAGATTAAGACAGACGATAC 613  
361 GCTGCTCCGAGAGCTTGGCCAGCCGATCAGAGACAAAGATGAGCAGAACGAGACC 420  
614 CCAAGTTTGTGCGCATGCTGGGCGGCTCTTCTGTGATGTTCTGGCAAGTTCAACT 673  
421 CCAAGTTTGTGCGCATGCTGGGCGGCTCTTCTGTGATGTTCTGGCAAGTTCAACT 480  
674 CTGCTGTGCTG 684

Db 481 CTGCTGTGCTG 491

RESULT 7  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS

BF602079 488 bp mRNA 'linear EST 25-APR-2001  
267224 MARC 3BOV Bos taurus cDNA 5', mRNA sequence.  
BF602079  
BF602079.1 GI:11699302  
EST.  
Bos taurus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovidae; Bovine; Bos.

1 (bases 1 to 488)  
Smith, T.P.L., Grose, W.M., Freking, B.A., Roberts, A.J., Stone, R.T., Casas, E., Wray, J.E., White, J., Cho, J., Fahrenkrug, S.C., Bennett, G.L., Heaton, M.P., Laegreid, W.W., Rohrer, G.A., Chitko-McKown, C.G., Perte, G., Holt, I., Karamycheva, S., Liang, F., Quackenbush, J., and Keeler, J.W.

Sequence evaluation of four pooled-tissue normalized bovine cDNA libraries and construction of a gene index for cattle

Genome Res. 11 (4), 626-630 (2001)

Contact: Smith TPL  
USDA, ARS, US Meat Animal Research Center  
PO Box 166, Clay Center, NE 68933-0166, USA  
Tel: 402 762 4366  
Fax: 402 762 4390  
Email: smith@email.marc.usda.gov

Single pass sequencing. Bases called and alt trimmed with phred v0.980904.e. Vector identified by cross\_match with the -minscore 18 and -mismatch 12 options.

PCR Primers  
FORWARD: AGGAACAGCTATGACCAT  
BACKWARD: GTTTCCTCAGTCAGCAGC  
Plate: 43 row: E column: 5  
Seq primer: ATTAGGTGACACTATAG.

FEATURES  
source  
1. .488  
/organism="Bos taurus"  
/db\_xref="taxon:9913"  
/clone\_1lb="MARC 3BOV"  
/tissue\_type="pooled"  
/lab\_host="DH10B"  
/note="Vector: PCMV SPORT6; Site\_1: NotI; Site\_2: SalI; Library made from pooled tissue from marrow, alveolar macrophage, ovary, fetal semitendinosus muscle, and fetal longissimus muscle."

BASE COUNT 97 a 137 c 135 g 119 t

ORIGIN

Query Match 23.0%; Score 288.2; DB 12; Length 488;  
Best Local Similarity 75.7%; Pred. No. 1.1e-70;  
Matches 370; Conservative 0; Mismatches 118; Indels 1; Gaps 1;

604 GCAAGATACCCAGTTGCTGTCATGCTGGCGCCCTCTTCTGTGATGTTCTGCGCA 663  
1 GCAAGAGCCCCAGTTGTTTCAACATGCTGGGCACCCCTCTTGTGATTAATTCGGCG 60  
664 AGTTCAACTGTGCTGCTGAGAGTCCATGCAAGGAAGATGCCGTTCACACC 723  
61 AGTTCAACTGTGCTGCTGAGAGTCCATGCAAGTAAGAA-GAACATGCGCGTGTTCACACC 119  
724 TACTATGCTAGAGATGAGCTGAGTGAAGCCATCTCAGGGTCACTCTTGGCTACCCC 783  
120 TACTACGCGCTGCGGCTTAGACAGTGAACGCGCATCTTAATGTCAGCCTTGGCTACCC 179  
784 CAAAGGAAGATCAAGAACTTAATGTCAGTGGGTGTGAGAGAGCGTGGCTGTG 843  
180 CAAAGGAAGATCAATATGATCTCAATCCACAGCAGTGTGGGAGGAGTGTGGCTGTG 239

```

OY      844  GGTACTCTCGTGCACCTGATCCCTTCCTCGTGGCTTGCCATGCTGGGCTTGTGGCT  903
Db      240  GGTGCCCTCTTCTTAACTGATCCATGCTCCTTGGCTTGGCATGTGCTGGGATCTGTGCT  299
OY      904  GGGTGATGTCCTCGTGGGGGAGCCAAAGTACCTGCCGGGGTGTGTGAACCGAGTCTGGGG  963
Db      300  GGGATGATGTCATCGGGGGGATTCAGATACCTGCAGGCGGTGTCTCCAGGAAGGTGCNA  359
OY      964  ATTCCCAAGCTCCATCATATGGGGTACCACTTCAGCTTGCTGGGTCTGTTGGAGAAGATC  1023
Db      360  CTCACGAACTTACGCGCGGTGCATCACCTTTGGCTTGGCGGGTCTGTTGGAGGAT  419
OY      1024  ATCTCATTTGCTGCTGCTGCTGCTTGTATACCGTGGAGCGGCAAGGCAATGATTTGGCTTC  1083
Db      420  GTCAACATGATGCTGATGTCGCGCTTCAAGCGCCAGGAGATGATGATCCAGCTCGGCTAC  479
OY      1084  CAGTCTCTC 1092
Db      480  AAGTACTC 488

RESULT 8
LOCUS   BG944752.c
DEFINITION BG944752 498 bp mRNA linear EST 11-JUN-2001
LOCUS   ax54d09.x1 Proliferating Human Erythroid Cells (LCB:ax library)
DEFINITION Homo sapiens cDNA clone ax54d09 random, mRNA sequence.
ACCESSION BG944752
VERSION   BG944752.1 GI:14344124
KEYWORDS  EST.
SOURCE   human.
ORGANISM Homo sapiens
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 498)
AUTHORS  Gubin,A.N., Njoroge,J.M., Bouffard,G.G. and Miller,J.L.
TITLE     Gene expression in proliferating human erythroid cells
JOURNAL   Genomics 59 (2), 168-177 (1999)
MEDLINE   99339981
COMMENT   Contact: Jeffery L. Miller
           Laboratory of Chemical Biology
           National Institute of Diabetes and Digestive and Kidney Diseases
           Building 10, Room 9B17, National Institutes of Health, Bethesda, MD
           20892, USA
           Tel: 301 402 2373
           Fax: 301 435 5148
           Email: jmf@nih.gov
           DNA Sequencing and analyses by National Institutes of Health
           Intramural Sequencing Center (NISC).
           Plate: 54 row: d column: 09
           Seq primer: -21m13 forward primer (ABI).
           Location/Qualifiers
             1..498
               /organism="Homo sapiens"
               /db_xref="taxon:9606"
               /clone="ax54d09"
               /clone_1ib="Proliferating Human Erythroid Cells (LCB:ax
               library)"
               /sex="unknown"
               /tissue_type="Blood"
               /cell_type="Erythroid Cells"
               /cell_line="Primary Culture of Peripheral Blood
               Mononuclear Cells"
               /dev_stage="Progenitor; EPO responsive CD71++++"
               /lab_host="SOLR"
               /note="Organ: Blood, Vector: Lambda ZMP II; Site 1: EcoRI
               Site 2: EcoRI; 65,000 proliferating erythroid cells from
               the buffy coat of a blood donation were obtained by flow
               cytometric separation after a 5-day culture period in the
               presence of erythropoietin. Total RNA was purified from
               the sorted cell population using TRIzol reagent. RNA (0.3
               ug) was converted into double stranded cDNA using
               Clontech's Capfinder cDNA Library Construction Kit

```

						(Clontech) according to the manufacturer's protocol and cloned into EcoRI digested lambda Zap II vector (Stratagene). The phage library was amplified once prior to in vivo excision in SOLR cells. Individual colonies were grown, and the cDNA inserts were sequenced in high throughput (NIH intramural sequencing center <a href="http://www.nisic.nih.gov/">http://www.nisic.nih.gov/</a> )."
BASE COUNT	134 a	99 c	106 g	159 t		
ORIGIN						
Query Match	21.6%;	Score 271.2;	DB 13;	Length 498;		
Best Local Similarity	94.0%;	Pred. No. 7.5e-66;				
Matches 282;	Conservative 0;	Mismatches 18;	Indels 0;	Gaps 0;		
OY	955	GTGCTGGGGATTCCCAACAGCTTCATCATGAGGCTRCACACTTGCAGTTGCTGGCTTCGCTT	1014			
Db	498	GTGCTGGGGATTCCACCAATTTCCGTCAATGACATCCATCCTTCAGTTGCTGGCTTCGCTT	439			
OY	1015	GGAGAGATCATCTACATTGTGTGTGTGGTGTAACCGTCGAGCCGGCAATGGCATG	1074			
Db	438	GGAGAGATCATCTACATTGTGTGTGTGGTGTAACCGTCGAGCCGGCAATGGCATG	379			
OY	1075	ATTGACTTCCAGAGTCCCTCCTCAGCATTTGGGGAACTCACCTTGGCCATGATAGCTTC	1134			
Db	378	ATTGACTTCCAGAGTCCCTCCTCAGCATTTGGGGAACTCACCTTGGCCATGATAGCTTC	319			
OY	1135	ACGTCGTCTCTCTGACAGGTTTGCTCTTAATCTTAAATRTGAANAACACTATGAG	1194			
Db	318	ACGTCGTCTCTCTGACAGGTTTGCTCTTAATCTTAAATRTGAANAACACTATGAG	259			
OY	1195	GCTAATATTTTGTGACCAAGTTTCTCGAAGTTTCTCCTCAATTTGGCTGTGGATTTTAA	1254			
Db	258	GCTAATATTTTGTGACCAAGTTTCTCGAAGTTTCTCCTCAATTTGGCTGTGGATTTTAA	199			
RESULT 9						
R10548						
LOCUS	R10548	416 bp	mRNA	linear	EST 06-APR-1995	
DEFINITION	yf1e01.r1 Soares fetal liver spleen INFLS Homo sapiens cDNA clone IMAGE:128472 5' similar to gb:X54534 BLOOD GROUP RH(D) POLYPEPTIDE (HUMAN); mRNA sequence.					
ACCESSION	R10548					
VERSION	R10548.1	GI:762504				
KEYWORDS	EST.					
SOURCE	human.					
ORGANISM	Homo sapiens					
JOURNAL	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.					
REFERENCE	1 (bases 1 to 416)					
AUTHORS	Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F., Trevasakis,E., Waterston,R., Williamson,A., Wohldmann,P. and Wilson,R.					
TITLE	The WashU-Merck EST Project					
COMMENT	Unpublished (1995) Contact: Wilson RK Washington University School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108 Tel: 314 286 1800 Fax: 314 286 1810 Email: est@watson.wustl.edu Insert Size: 1680 High quality sequence stops: 202 Source: IMAGE Consortium, LNL This clone is available royalty-free through LNL; contact the IMAGE Consortium ( <a href="mailto:info@image.jlnl.gov">info@image.jlnl.gov</a> ) for further information. Insert Length: 1680 Std Error: 0.00 Seq primer: ML3RP1 High quality sequence stop: 202. Location/Qualifiers 1..416 /organism="Homo sapiens" /db_xref="GDB:1480633" /db_xref="taxon:9606"					
FEATURES						
source						

Query Match	Best Local Similarity	Matches	315; Conservative	21.6%; Score 270.4; DB 14; Length 416; 97.2%; Pred. No. 1.2e-65; 0; Mismatches 5; Indels 4; Gaps 4;
QY	1	ATGAGCTCTAAGTACCCGCGGTCTGTCCGGCGCTGCTGCCCTCTGAGCCCTTAACACTG	60	
Db	46	ATGAGCTCTAAGTACCCGCGGTCTGTCCGGCGCTGCTGCCCTCTGAGCCCTTAACACTG	105	
QY	61	GAGGAGCTCTCATCTCTCTTTATTTTTTTTAAACCAATGACGCTTCTTAAAGAT	120	
Db	106	GAGGAGCTCTCATCTCTCTTTATTTTTTTTAAACCAATGACGCTTCTTAAAGAT	165	
QY	121	CAAAAGGGGCTGTGGCATTATCAAGTTGGCCAAAGATCCGTGATGGCGGCCATT	180	
Db	166	CAAAAGGGGCTGTGGCATTATCAAGTTGGCCAAAGATCCGTGATGGCGGCCATT	225	
QY	481	GCGTTGGGCTTCCCTCACTCGAGTTTCCGGAACACAGCTGAGAGAGTGTGGCTTCAAC	240	
Db	226	GCGTTGGGCTTCCCTCACTCGAGTTTCCGGAACACAGCTGAGAGAGTGTGGCTTCAAC	285	
QY	241	CTCTTCATGCT-GGGCGTTGGTGTGAGT-GGGCAATCTGCT-GGACGGCTTCTGAGC	297	
Db	286	CTCTTCATGCTGGGCGCTTGGTGTGAGTGGGGCAATCTGCTGGGAGCGCTTCTGAGC	345	
QY	298	CAG-TTCCCTTCTGGGAAGTGT	320	
Db	346	CAGTTCCCTTCTGGGAAGTGT	369	
RESULT 10				
LOCUS	TR84327			
DEFINITION	Y3d37a05.r1 Soares fetal liver spleen INFLS Homo sapiens cDNA clone	343 bp	mRNA	linear EST 16-MAR-1995
ACCESSION	Y3d37a05	110384	5'	similar to gb:X54534 BLOOD GROUP RH(D) POLYPEPTIDE (HUMAN); mRNA sequence.
VERSION	TR84327.1	GI:712615		
KEYWORDS	EST.			
SOURCE	human.			
ORGANISM	Homo sapiens			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
AUTHORS	1 (baes 1 to 343)			
JOURNAL	Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M., Holman, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M., Parsons, J., Rifkin, L., Rohlfing, T., Soares, M., Tan, F., Trevisan, E., Waterston, R., Williamson, A., Woldmann, P. and Wilson, R.			
COMMENT	The WashU-Merck EST Project			
	Unpublished (1995)			
	Contact: Wilson RK			
	Washington University School of Medicine			
	4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108			
	Tel: 314 286 1800			
	Fax: 314 286 1810			
	Email: est@watson.wustl.edu			
	Insert Size: 1034			

<hr/>					
High quality sequence stops: 295 Source: IMAGE Consortium, LNLN This clone is available royalty-free through LNLN; contact the IMAGE Consortium ( <a href="mailto:info@image.lnl.gov">info@image.lnl.gov</a> ) for further information. Insert Length: 1034 Std Error: 0.00 Seq primer: M13RP1  High quality sequence stop: 295.					
Location/Qualifiers					
1..343					
/organism="Homo sapiens"					
/db_xref="GDB:466001"					
/db_xref="taxon:9606"					
/clone="IMAGE:110384"					
/clone_lib="Soares fetal liver spleen INFLS"					
/sex="male"					
/dev_stage="20 week-post conception fetus"					
/lab_host="DH10B (ampicillin resistant)"					
/note="Organ: Liver and Spleen; Vector: pTR73D (Pharmacia)					
with a modified polylinker; Site_1: Pac I; Site_2: Eco RI;					
1st strand cDNA was ligated with a Pac I - oligo(dT) primer					
[5'. AACTGGAGAATTAATTAAAGATCTTTTCTTTTTTTTGT 3'] ,					
double-stranded cDNA was ligated to Eco RI adaptors					
(Pharmacia), digested with Pac I and cloned into the Pac I					
and Eco RI sites of the modified pTR73 vector. Library					
went through one round of normalization. Library					
constructed by Bento Soares and M.Facima Bonaldo."					
<hr/>					
BASE COUNT	78 a	88 c	92 g	85 t	
ORIGIN					
<hr/>					
Query Match	20.2%	Score 253.6;	DB 14;	Length 343;	
Best Local Similarity	94.6%;	Pred. No. 6.4e-61;	Mismatches 14;	Indels 3;	Gaps 3;
Matches 295;	Conservative 0;	Mismatches 14;	Indels 3;	Gaps 3;	
OY	497	AACATGAACATGATGCATCTACGTGTCGCGAGCCATAATTTTGAGCTGCTGCGCT-GG	555		
Db	1	ACATGAACCTTAGAGGACTTACGCGITTCGAGCCTAATTTGGGCTGACGTGGCTGGG	60		
OY	556	TGCCCTGCCAAAAGCCTCTATCCCAGAGGAAACGAGATTAAGATCAACAGCACACGATACC	615		
Db	61	TGCCCTGCCAAAAGCCTCTATCCCAGAGGAAACGAGATTAAGATCAACAGCACACGATACC	120		
OY	616	AGTTGTCGTCATGCTGGGGGCGCCCTCTCTTGAGCATGCTGGCCAAGTTCAACTCT	675		
Db	121	AGTTGTCGTCATGCTGGGGGCGCCCTCTCTTGAGCATGCTGGCCAAGTTCAACTCT	180		
OY	676	GCTGCTGAGAGAAGTCATCGAAAGAAAGATGCCGTTCACAACACCTACTATGCTGTA	735		
Db	181	GCTGCTGAGAGAAGTCATCGAAAGAAAGATGCCGTTCACAACACCTACTATGCTGTA	240		
OY	736	GCAAGTCAGCGT-GGTGACAGCGCATCTCAGGGTCATCTCT--GGCTCACCCCCAACAGGAGA	793		
Db	241	GCAAGTCAGGTGGGGGACAGCGCATCTCAGGGTCATCTCTGGGCTCACCCCCAACAGGAGA	300		
OY	794	TCAGCAACACTT 805			
Db	301	TCAGCAACACTT 312			
<hr/>					
RESULT 11					
BFI91605					
LOCUS	BFI91605	515 bp	mRNA	linear	EST 02-NOV-2000
DEFINITION	239237 MARC 2P1G Sus scrofa cDNA 5', mRNA sequence.				
ACCESSION	BF191605				
VERSION	BF191605.1 GI:11074974				
KEYWORDS	EST.				
SOURCE					
ORGANISM	Sus scrofa				
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;					
Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.					
1 (bases 1 to 515)					
Fahrenkrug,S.C., Freking,B.A., Rohrer,G.A., Smith,T.P.L., Casas,E.,					
Stone,R.T., Heaton,M.P., Grosse,W.M., Bennett,G.A., Laegreid,W.W.					
and Keeler,J.W.					
Design and use of two pooled tissue normalized cDNA libraries for					

JOURNAL  
COMMENT

EST discovery in swine  
Unpublished (2000)  
Contact: Smith TPL  
USDA, ARS, US Meat Animal Research Center  
PO Box 166, Clay Center, NE 68933-0166, USA  
Tel: 402 762 4366  
Fax: 402 762 4390  
Email: smtlth@mail.marc.usda.gov  
Single pass sequencing. Bases called and alt trimmed with phred  
v0.980904.e. Vector identified by cross\_match with the -minscore 18  
and -minmatch 12 options.

PCR Primers  
FORWARD: AGGAACAGCTATGACCAT  
BACKWARD: GTTTCACGTCACGACG  
Plate: 67 row: A column: 8  
Seq primer: ATTAGGAGACATATAG.

## FEATURES

## Source

Location/Qualifiers  
1. .515  
/organism="Sus scrofa"  
/db\_xref="taxon:9823"  
/clone\_lib="MARC 2P1G"  
/tissue\_type="pooled"  
/lab\_host="DH10B"  
/note="Vector: pCMV SPORT6; Site\_1: NotI; Site\_2: SalI;  
Library made from pooled tissue from testis, ovary,  
endometrium, hypothalamus, pituitary, and placenta."  
BASE COUNT 103 a 143 c 148 g 121 t  
ORIGIN

Query Match 19.4%; Score 243; DB 12; Length 515;  
Best Local Similarity 73.1%; Pred. No. 7.8e-58;  
Matches 312; Conservative 0; Mismatches 115; Indels 0; Gaps 0;

136 GCATCTTATCAAGTTCGGCAAGATCTGAACCGTGAGGCGGCATATGCTGGGCTTCTC 195  
Db . 89 GCGAGCGACCGGCTTCAGGATGTGTATCATATGAGCCCTTGCGCTTGCTCTC 148  
QY 196 ACCTGAGCTTCGCGAGACACAGCTGAGCAAGTGTGCGCTTCAACCTTCAATGCTGGCG 255  
Db 149 AATAATCTTTGGCGAGACACTGCGAGCAAGCGGCTTCAATCTTTCTGCTGGCT 208  
QY 256 CTGTGTGTGAGTGGGCAATCTCTGTGACGCGCTTCTGAGCCAGTCTCCCTTGGGAG 315  
Db 209 CTCGGGGTACAGTTGACCTCTCTGTGATGCTTCTGAGCCGACCTTTCCAGGAG 268  
QY 316 GTGGTCATCAACTGTAGTATTGGCGTGGCCACATGAGTGTGTCGCTGCTGATC 375  
Db 269 ATGAGCATCAATCTGACTAGTATTGAGAGCGCACCGTGAAGCTGCTGCTGATC 328  
QY 376 TCAGTGAATGCTGTCTTGGGGAAGGTCAACTTGGCGCAGTTGTGATGCTGCTG 435  
Db 329 TCAAGGGGTGCATCTCTGGGAAGGTCAACCTGTGCACTGTGTTTATGACATGATA 388  
QY 436 GAGGTGACAGCTTTAGGCAACTGAGATGTGATCAGTAAATATCTTCAACACAGACTAC 495  
Db 389 GAGGTGACCGCTTACAGACCAAGAGATGTGCGCAAGATTCCTCAATGTGAGAAC 448  
QY 446 CACATGAACATGATCAATCTACGTCTTCCGAGGCTATTTGGGCTGTCTGTGGCGCTGG 555  
Db 449 CACGTAACATGATGATCAATCATGATGTGCGGCTATTTCGGGCTGACGTGTGAGCTGC 508  
QY 556 TGCCTGC 562  
Db 509 TGCCTGC 515

RESULT 12  
AA259801

LOCUS AA259801 491 bp mRNA linear EST 18-MAR-1997  
DEFINITION var7g11.t1 Soares mouse 3ME12.5 Mus musculus cDNA clone  
IMAGE:746468 5' similar to gb:X54534 BLOOD GROUP RH(D) POLYPEPTIDE  
(HUMAN); mRNA sequence.  
ACCESSION AA259801

VERSION  
KEYWORDS  
SOURCE  
ORGANISMREFERENCE  
AUTHORSTITLE  
JOURNAL  
COMMENT

AA259801.1 GI:1896287  
EST.  
house mouse.  
Mus musculus  
Mammalia, Metazoa: Chordata; Craniata; Vertebrata; Euteleostomi;  
Eumaryota; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 491)  
Marrs, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,  
Gaisel, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M.,  
Schellenberg, K., Stepoe, M., Tan, F., Underwood, K., Moore, B.,  
Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and  
Waterston, R.  
The WashU-HMI Mouse EST Project  
Unpublished (1996)  
Contact: Marra M/Mouse EST Project.  
WashU-HMI Mouse EST Project  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: mouseest@watson.wustl.edu  
This clone is available royalty-free through LNL; contact the  
IMAGE Consortium (info@image.lnl.gov) for further information.  
MW:455452  
Seq primer: -28ml3 rev2 ET from Amersham  
High quality sequence stop: 481.  
Location/Qualifiers  
1. .491  
/organism="Mus musculus"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/clone\_lib="IMAGE:746468"  
/clone\_lib="Soares mouse 3ME12.5"  
/sex="unknown"  
/tissue\_type="fetus"  
/dev\_stage="12.5dpc total fetus"  
/lab\_host="DH10B"  
/note="Organ: whole fetus; Vector: pT73D-Bac (Pharmacia).  
with a modified polylinker; Site\_1: Not I; Site\_2: Eco RI;  
1st strand cDNA was primed with a Not I - oligo(dT) primer  
[5', TGTTACCAATCGAAGTGGAGCGCGGCTTATTTTGTGTTT  
3'], on total mouse RNA [provided by Minoru Ko, Wayne  
State Univ.]; double-stranded cDNA was ligated to Eco RI  
adaptors (Pharmacia), digested with Not I and cloned into  
the Not I and Eco RI sites of the modified pT73 vector.  
Library went through one round of normalization, and was  
constructed by Bento Soares and M. Fatima Bonaldo."

FEATURES  
Source

BASE COUNT 97 a 145 c 127 g 122 t  
ORIGIN  
Query Match 18.9%; Score 236.8; DB 9; Length 491;  
Best Local Similarity 71.4%; Pred. No. 4.4e-56;  
Matches 354; Conservative 0; Mismatches 137; Indels 5; Gaps 3;

QY 16 CCGGGTCTGTCCGCGCTGCTGCCCCCTGAGCCCTTAACACTGGAAGCCTTCATT 75  
Db 1 CCAAGGTCCCTCCGCTGCTGCTGCCCCCTG- GGCCTTGAGCTACAGACAGCTTTAGT 59  
QY 76 CTCTCTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 135  
Db 60 CTCTCTTTCTTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 116  
QY 136 GCATCTTATCAAGTTCGGCAAGATCTGACCGTGAATGGCGGCATATGCTGGGCTTCTC 195  
Db 117 GCGAGCTATCAAGTTCCTCGGAATTTGACCTCATGGAAGC- TTGGGTTGGGCTTCTG 175  
QY 196 ACCTGAGCTTCGCGAGACACAGCTGAGCAAGTGTGCGCTTCAACCTTTCATGCTGGCG 255  
Db 176 TCTCTGTCTTTCCGAGACACAGCTGAGCAAGTGTGCGCTTCAACCTTTCATGCTGGCG 235  
QY 256 CTGTGTGTGAGTGGGCAATCTCTGTGACGCGCTTCTGAGCCAGTTCCTTGGGAG 315  
Db 236 CTCGGGGTGCAGGGAACATCTTCTGTGACCATTTCTGCGGCGCAGGCTCTCAATGGAAC 295

FEATURES	source	Location/Qualifiers
Db	296	AAAGTCAACAACTCTGTCCAGCATCCAGTACGATACAGACACTTACCTGTGTATC 355
Qy	316	GTGTGATCAACACTGTTACGATTTGGCTGGCCACACATGAGGCTTGTGGTGTATC 375
Db	356	TCACGGGGCGCTGTCCTGGGAAAGGTCAACTGTGGCAGCTGACCGTATGGTGTATG 415
Qy	436	GAGGTGACAGCTTTAGGCAACTGAGATGTCATCATTAATTTTCAACACAGACTAC 495
Db	416	GAGGCAATGCGCTTTGGTGGCCATCATGATTTGGCCGACGAAAGGTCTTCAAATGACAGAA 475
Qy	496	CACATGAACATGATGC 511
Db	476	CACATCATCATGATGC 491
RESULT 13		
LOCUS	BB847808	446 bp mRNA linear EST 26-NOV-2001
DEFINITION	BB847808	CDNA clone F530107A15 5', mRNA sequence.
ACCESSION	BB847808	
VERSION	BB847808.1	GI:17086183
KEYWORDS		EST.
SOURCE		house mouse.
ORGANISM		Mus musculus
REFERENCE		Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eultheria; Rodentia; Sciurognathi; Muridae; Mus.
AUTHORS		1 (bases 1 to 446) Akimura,T., Arakawa,T., Carninci,P., Furuno,M., Hanagaki,T., Hayatsu,N., Hiramoto,K., Hiroka,T., Hirozane,T., Imotani,K., Ishii,Y., Ito,M., Kawai,J., Kojima,Y., Kono,H., Kouda,M., Matsuyama,T., Nakamura,M., Nishi,K., Nomura,K., Numasaki,R., Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K., Sakazume,N., Sasaki,D., Sato,K., Shibata,K., Shitagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H., Tagawa,A., Takahashi,F., Takeku-Akahori,S., Tanaka,T., Tomaru,A., Toyota,T., Wachihi,A., Yasunishi,A., Muramatsu,M. and Hayashizaki,Y.
TITLE		RIKEN Encyclopedia of Mouse Full-length cDNAs (Akimura,T., et al. 2001)
JOURNAL		Unpublished (2001)
COMMENT		Contact: Yoshihide Hayashizaki Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center(GSC), Yokohama Institute The Institute of Physical and Chemical Research (RIKEN) 1-7-22 Suenho-cho, Tsurumi-Ku, Yokohama, Kanagawa 230-0045, Japan Tel: 81-45-503-9222 Fax: 81-45-503-9216 Email: genome-res@gsc.riken.go.jp, URL: http://genome.gsc.riken.go.jp/ Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Kono,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y. Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000) wagi,K., Fujiwaka,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsura,S., Kawai,U., Okazaki,Y., Muramatsu,M., Inoue,Y., Kita,A. and Hayashizaki,Y. RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000) Kono,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P., Sugahara,Y. and Hayashizaki,Y. Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001) please visit our web site (http://genome.gsc.riken.go.jp) for further details. e mouse tissues.
		1..446

```

/organism="Mus musculus"
/db xref="taxon:10090"
/clone="F530107A15"
/clone_11b="Riken full-length enriched, adult male kidney"
/sex="male"
/tissue_type="kidney"
/dev stage="adult"
/lab host="SOLR"
/notes="Site 1: XhoI, Site 2: SstI; cDNA library was
prepared and sequenced in Mouse Genome Encyclopedia
Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in
RIKEN. Division of Experimental Animal Research in Riken
contributed to prepare mouse tissues. 1st strand cDNA was
primed with a primer [5']
GAAGAGAGAGCGCGCCGCACTCGAGTTTCTTTTCTTTTAA 3'], cDNA was
prepared by using triethanolamine-activated reverse
transcriptase and subsequently enriched for full-length by
cap-trapper. Second strand cDNA was prepared with the
primer adapter of sequence [5']
GAAGAGAGAGGATCCAGAGCTCAATTAATTAATTAACCCCCCCCC 3'].
cDNA was cleaved with XhoI and SstI. "

BASE COUNT      87 a      139 c      116 g      104 t
ORIGIN
Query Match      18.5%; Score 231.8; DB 10; Length 446;
Best Local Similarity 74.5%; Pred. No.1,1e-54;
Matches 306; Conservative 0; Mismatches 102; Indels 3; Gaps 1;

QY      1  ATGAGCTCTAAGTACCGCGGTCTGTGCGGCGTGCCTGCTCTGTGGGCCCTTAACGTG 60
Dy      39  ATGGGCTCTAAGTACCAACGCGTCCCTCGCGTGCCTGCGCCCTGTGGCCCTTGAAGCTA 98
QY      61  GAAGCAGCTCTCATCTTCCTCTTCTAATTTTAAACCACTATGACGCTTCTTAGAGGAT 120
Dy      99  CAGACAGCTTTTAAGTCTCTCTCTTCTGTTTTCATCCCCACGACACAGCCAGGTGAT 158
QY      121  CAAAGGGGCTCGTGGCATCTTATCAAGTTGGCCAAATCTGACCGTGATGGCGGCATT 180
Dy      159  CACA---GGTTCATGGGAGAGCTATCAAGTCTCTCGGAATTGACCTTCATGGACGCTTG 215
QY      181  GGCCTTGGGCTTCTCACTCTGAGTTTCCGAGACACAGCTGGAGCAGTGGCCTTCAAC 240
Dy      216  GGCCTTGGGCTTCTCTGCTGCTGCTTTCGAGACACAGCTGGAGCAGTGGCCTTCAAC 275
QY      241  CTTCTTACAGCTGGGCTGTGTGTCAGATGGGCAATCCGCTGAGAGCGCTTCTGAGCCAG 300
Dy      276  CTCCTTCAATTTGGCCCTCGGGGTGCAAGGAACAATCTTGTCTGAGACATTTCTGGCCAG 335
QY      301  TTCCCTTCTGGGAAGGTGTGTCATCACACTGTTCAATTTGGCTGGCCACCATAGTGTCT 360
Dy      336  GTCTCCCATATGGAACAAGATCAACAATCTGTCCAGATCCAGATAGCTACCATGAGCACC 395
QY      361  TTGTGGGCTGATCTCAGTGGATGCTGTCTTGGGGAAGGTCACTTGGGCG 411
Dy      396  TTACCTGTGCTGATCTCAGCGGCGCTGTCTTGGGGAAGGTCAACTGTGTG 446

RESULT 14
LOCUS      BG087038              454 bp      mRNA      linear      EST 26-JAN-2001
DEFINITION H3134C09-5 N1A Mouse 15K cDNA Clone Set Mus musculus cDNA clone
ACCESSION  BG087038
VERSION    BG087038.1 GI:12569613
KEYWORDS   EST.
SOURCE     house mouse.
ORGANISM  Mus musculus.
Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathu; Muridae; Murinae; Mus.
1 (bases 1 to 454)
Kargul,G.J., Dudekula,D.B., Qian,Y., Lim,M.K., Varadaraj,S.A., Tanaka
,T.S., Carter,M.G. and Ko,M.S.H.

```







THIS PAGE BLANK (USPTO)

GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 7, 2003, 18:42:20 ; Search time 3621 Seconds

(without alignments)  
10078.685 Million cell updates/sec

Title: US-09-600-714-41

Perfect score: 1254  
Sequence: 1 atgagctcctaagtaaccgcg.....attggctgttgattttaa 1254

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-Processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database :

GeneBml:\*  
1: gb\_da:\*  
2: gb\_htg:\*  
3: gb\_in:\*  
4: gb\_in:\*  
5: gb\_ov:\*  
6: gb\_pac:\*  
7: gb\_ph:\*  
8: gb\_pl:\*  
9: gb\_pr:\*  
10: gb\_ro:\*  
11: gb\_sts:\*  
12: gb\_sy:\*  
13: gb\_un:\*  
14: gb\_vi:\*  
15: em\_da:\*  
16: em\_fun:\*  
17: em\_hum:\*  
18: em\_in:\*  
19: em\_mu:\*  
20: em\_om:\*  
21: em\_or:\*  
22: em\_ov:\*  
23: em\_pac:\*  
24: em\_ph:\*  
25: em\_pl:\*  
26: em\_ro:\*  
27: em\_sts:\*  
28: em\_un:\*  
29: em\_vi:\*  
30: em\_htg\_hum:\*  
31: em\_htg\_inv:\*  
32: em\_htg\_other:\*  
33: em\_htg\_mus:\*  
34: em\_htg\_pln:\*  
35: em\_htg\_rnd:\*  
36: em\_htg\_man:\*  
37: em\_htg\_vit:\*  
38: em\_sy:\*  
39: em\_htgo\_hum:\*  
40: em\_htgo\_mus:\*  
41: em\_htgo\_other:\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1254	100.0	1254	6 AX022514	AX022514 Sequence
2	1252.4	99.9	1254	9 AB018969	AB018969 Homo sapi
3	1252.4	99.9	1354	9 HUMRHDPNTI	108429 Human RHD b
4	1252.4	99.9	1545	9 HSRHIT	X63094 H. sapiens m
5	1252.4	99.9	2790	9 HSRHIT	X63097 H. sapiens m
6	1250.8	99.7	1254	9 AB018966	AB018966 Homo sapi
7	1250.8	99.7	1254	9 AB018967	AB018967 Homo sapi
8	1247.6	99.5	1260	9 S78509	S78509 Rhd [human,
9	1246	99.4	1254	9 AB018968	AB018968 Homo sapi
10	1246	99.4	1254	9 S82449	S82449 Homo sapien
11	1242.8	99.1	1458	9 AF187846	AF187846 Homo sapi
12	1239.8	98.9	1251	9 S57971	S57971 Rh Polypept
13	1239.8	98.9	1254	9 AB046420	AB046420 Homo sapi
14	1238	98.7	1338	9 AB049754	AB049754 Homo sapi
15	1231.4	98.2	1247	9 AF037626	AF037626 Homo sapi
16	1226.8	97.8	1254	9 AB037270	AB037270 Homo sapi
17	1225.2	97.7	1336	9 AF312679	AF312679 Homo sapi
18	1222	97.4	1330	9 HS297026	HS297026 Homo sapien
19	1220.4	95.7	1254	9 AF177938	AF177938 Homo sapi
20	1199.6	95.7	1338	9 AB030388	AB030388 Homo sapi
21	1199.6	95.7	1338	9 AB049753	AB049753 Homo sapi
22	1191.6	95.0	1350	9 GORRHB	137053 Gorilla gor
23	1188.4	94.8	1358	9 AB018644	AB018644 Homo sapi
24	1186.8	94.6	1350	9 CHPRHC	137050 Pan troglod
25	1185.2	94.5	1384	6 HUMRHDPAA	189819 Sequence 1
26	1185.2	94.5	1384	9 HUMRHDPAA	M34015 Human Rh po
27	1185.2	94.5	1471	9 HSRH30A	X54534 Human mRNA
28	1183.6	94.4	1328	9 AB018645	AB018645 Homo sapi
29	1183.6	94.4	1350	9 CHPRHB	137049 Pan troglod
30	1182	94.3	1466	6 189820	189820 Sequence 2
31	1180.8	94.2	1427	9 AF056965	AF056965 Homo sapi
32	1180.6	94.1	1251	9 S57967	S57967 Rh polypept
33	1180.4	94.1	1350	9 CHPRHA	137048 Pan troglod
34	1169.2	93.2	1350	9 GORRHA	137052 Gorilla gor
35	1102.2	87.9	1270	9 AF012425	AF012425 Pongo pyg
36	1055.6	84.2	1350	9 MACRHA	137054 Macaca fasc
37	1049.2	83.7	1369	9 S70343	S70343 Rh-like [ma
38	1038	82.8	1269	9 AF012426	AF012426 Papio ham
39	972.8	77.6	1263	9 AF012429	AF012429 Callithrix
40	972	77.5	1263	9 AF012428	AF012428 Saimiri s
41	962.8	76.8	1216	9 AF0101479	AF0101479 Cebus ape
42	954.2	75.1	1209	9 AF102876	AF102876 Cebus ape
43	952	75.9	1216	9 AF052588	AF052588 Callithrix
44	941.2	75.1	951	9 S73913	S73913 Rh D-RhK562
45	939	74.9	966	9 S70174	S70174 Rhd-Rh Dloo

#### ALIGNMENTS

RESULT 1  
AX022514  
LOCUS AX022514  
DEFINITION Sequence 41 from Patent WO9337763.  
ACCESSION AX022514  
VERSION AX022514.1 GI:10046112  
KEYWORDS  
SOURCE unclassified.  
ORGANISM unclassified.  
REFERENCE 1 (bases 1 to 1254)  
AUTHORS Riegel, W.A. and Wagner, F.F.  
TITLE Novel nucleic acid molecules correlated with the rhesus weak d phenotype  
JOURNAL Patent: WO 9337763-A 41 29-JUL-1993;

FEELGEL WILLY A (DE) ; WAGNER FRANZ F (DE) ; DRK BLUTSPENDEDIENST  
BADEN WUE (DE)  
Location/Qualifiers  
1. .1254  
/organism="unidentified"  
/db\_xref="taxon:32644"  
1. .1254  
/note="unnamed protein product"

## CDS

/codon\_start=1  
/protein\_id="CAC07879.1"  
/db\_xref="GI:10946113"  
/translation="MSSKTPRSVRRCIPNALITLALILFFFTHYDASLEDOKL  
VASYQVQDLTVMAIGLPLTSSFRHSMSVAFLFMALVQMAILLDGLISQPL  
SGKAVITLFSRLTMSGLSVLSVDVLAQVLVWLVVLAIGLIRVINSI  
FNIDYHMMHIIYFAAFVGLSVAMCLPKPLPESTERDQATIPSLAMGLFLM  
FWPFSNALLRSPERKQAVENVYAVSVTAHSSSLAHPOGKISKTYHSAVLA  
GVAVGTSCHLIPSPMLAMVLGLVAGLISVQAGAYLGGCCNRUGIPIHSSIMGNFSL  
LGILGELIYIVLVLDVVGAGNMGIFQVLIISLESLAIVIALTSLGLTLNLT  
WKAPHEKYPDDVFMKFPHLAVGF"

BASE COUNT 241 a 322 c 351 g 340 t  
ORIGIN  
Query Match 100.0%; Score 1254; DB 6; Length 1254;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1254; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAGCTTAAGTACCCGCGGTCTGTCGGCGGCTGCTGCGCCCTTGAGCCCTTACACATG 60  
Db 1 ATGAGCTTAAGTACCCGCGGTCTGTCGGCGGCTGCTGCGCCCTTGAGCCCTTACACATG 60  
QY 61 GAAGAGCTCTCATCT 120  
Db 61 GAAGAGCTCTCATCT 120  
QY 121 CAAAGGGGCTCTGCGGCT 180  
Db 121 CAAAGGGGCTCTGCGGCT 180  
QY 181 GGCCTTGAGCT 240  
Db 181 GGCCTTGAGCT 240  
QY 241 CTCTTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300  
Db 241 CTCTTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300  
QY 301 TTCCCTTTTGGGAAGTGTATCATCACTGTTCACTATTTCCGCTGCGCCACCATGAGTGTCT 360  
Db 301 TTCCCTTTTGGGAAGTGTATCATCACTGTTCACTATTTCCGCTGCGCCACCATGAGTGTCT 360  
QY 361 TTGTCGCTGCTGATCTCACTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420  
Db 361 TTGTCGCTGCTGATCTCACTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420  
QY 421 GTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480  
Db 421 GTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480  
QY 481 TTCAACACAGCTACCAATGAACATGATGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540  
Db 481 TTCAACACAGCTACCAATGAACATGATGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540  
QY 541 CTGTCGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600  
Db 541 CTGTCGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600  
QY 601 ACAGCAACGATACCCAGTTTGTCTGCAATGCTGCGCGCCCTTCTTGTGATGTTCTGCG 660  
Db 601 ACAGCAACGATACCCAGTTTGTCTGCAATGCTGCGCGCCCTTCTTGTGATGTTCTGCG 660  
QY 661 CCAAGTTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720  
Db 661 CCAAGTTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720

QY 721 ACTTACTATGCTGTAGAGTACAGTCAAGTGTGACAGCCATCTCAGGGTCACTCTTGGCTCAC 780  
Db 721 ACTTACTATGCTGTAGAGTACAGTCAAGTGTGACAGCCATCTCAGGGTCACTCTTGGCTCAC 780  
QY 781 CCCCAGGGAAGATCAGACAGACTTATATGTGACAGTGCCTGCTTGGCAGAGGCGTGGCT 840  
Db 781 CCCCAGGGAAGATCAGACAGACTTATATGTGACAGTGCCTGCTTGGCAGAGGCGTGGCT 840  
QY 841 GTGGGTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900  
Db 841 GTGGGTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900  
QY 901 GCTGGGCTGATCTCTCGTGGGGAGCCAGAGTACCTGCGGGGTGTGTAACCGAGTCTG 960  
Db 901 GCTGGGCTGATCTCTCGTGGGGAGCCAGAGTACCTGCGGGGTGTGTAACCGAGTCTG 960  
QY 961 GGGATTCGCCACAGCTCATATATGGGCTACACACTTCAAGCTTGTGCTGCTGCTGAGAG 1020  
Db 961 GGGATTCGCCACAGCTCATATATGGGCTACACACTTCAAGCTTGTGCTGCTGCTGAGAG 1020  
QY 1021 ATCATCTACATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1080  
Db 1021 ATCATCTACATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1080  
QY 1081 TTCCAGGCTCTCTCAGACATTTGGGGAAGTCAAGCTTGGCCATCTGATAGCTCTCAGCTCT 1140  
Db 1081 TTCCAGGCTCTCTCAGACATTTGGGGAAGTCAAGCTTGGCCATCTGATAGCTCTCAGCTCT 1140  
QY 1141 GGTCTCTGACAGGTTGCTCTTAAATCTTAAATATGGAAGACCTCATGAGGCTTAA 1200  
Db 1141 GGTCTCTGACAGGTTGCTCTTAAATCTTAAATATGGAAGACCTCATGAGGCTTAA 1200  
QY 1201 TATTTGATGACCAAGTTTCTGGAAGTTTCTCATTTGGCTGTGATTTAA 1254  
Db 1201 TATTTGATGACCAAGTTTCTGGAAGTTTCTCATTTGGCTGTGATTTAA 1254

RESULT 2  
ABO18969  
LOCUS  
DEFINITION  
Homo sapiens RHDY0 mRNA for Rh blood group D antigen (RHD),  
complete cds.  
ACCESSION  
ABO18969  
VERSION  
ABO18969.1 GI:5360898  
KEYWORDS  
RHDY0; Rh blood group D antigen (RHD).  
SOURCE  
Homo sapiens (isolate:YO) cDNA to mRNA.  
ORGANISM  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE  
AUTHORS  
Hyodo, H., Ishikawa, Y., Kashiwase, K., Ogawa, A., Watanabe, Y.,  
Tsuneyama, H., Toyoda, C., Uchikawa, M., Akaza, T. and Fujii, T.  
TITLE  
Polymorphisms of RHDY0 in Japanese  
Unpublished  
REFERENCE  
AUTHORS  
Uchikawa, M., Hyodo, H. and Ishikawa, Y.  
TITLE  
Submitted (20-OCT-1998) Hironobu Hyodo, Japanese Red Cross Central  
Blood Center, Research; 4-1-31, Hiroo, Shibuya, Tokyo 150-0012,  
Japan (E-mail:hyodo@hla.cbc.jrc.or.jp, Tel:+81-3-5485-6009,  
Fax:+81-3-3406-7892)

FEATURES  
source  
Location/Qualifiers  
1. .1254  
/organism="Homo sapiens"  
/isolate="YO"  
/db\_xref="taxon:9606"  
1. .1254  
/gene="RHDY0"  
1. .1254  
/gene="RHDY0"  
/note="Rhd variant RhdY0"  
/codon\_start=1

```

/product="Rh blood group D antigen (RHD)"
/db_xref="GI:5360839"
/translation="MSSKYPRSYRCLPLMALTLLEAALLLPFYTHYASLEDOGL
VASVYQDPLTWAAIGLGLFSSPRRHSMSVAFLPMLALGVMAIILIDGLSOP
SGRVVITLPSIRLATMSALSVLSVDAVIGKVALQLVVWVLEVTALGNLRVINSI
FNTDYHNMNMHIYVFAVFGLSVAMCLPEFLPEGTBDKQATATPBLISANLGLFLMM
FMSFNSALIRSPDKRKNAVFNTRYAVAVSVVAISGSSLAHQGKISKYVSAVLA
GVAVGTSCHLIPSPMLAMVLGVLGAGISVGAKYLPGCCNRVLGI.PHSIMGNFSL
LGLLGEIYIVLIVLDTVGAGNMIGFOVLSIGELSLAIVIALTSGLLTGLLNLKI
WKAPRAKYPDDQVFWKPRHLAVGF"
BASE COUNT      242 a      322 c      350 g      340 t
ORIGIN

Query Match      99.9%; Score 1252.4; DB 9; Length 1254;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1251; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 ATGAGCTCTAAGTACCGCGGCTGTGTCGCGCTGCTGCCCTCTGCGCCCTTAACATG 60
DB      1 ATGAGCTCTAAGTACCGCGGCTGTGTCGCGCTGCTGCCCTCTGCGCCCTTAACATG 60

QY      61 GAAGGAGCTCATATTCCTCTCTATTTTTCACCAATAGACGCTTCTTGAGAGAT 120
DB      61 GAAGGAGCTCATATTCCTCTCTATTTTTCACCAATAGACGCTTCTTGAGAGAT 120

QY      121 CAAAGGGGCTCGTGGAGCTTATCAAGTTGGCCAAATGTGACCGTGATGGCGGCATT 180
DB      121 CAAAGGGGCTCGTGGAGCTTATCAAGTTGGCCAAATGTGACCGTGATGGCGGCATT 180

QY      121 CAAAGGGGCTCGTGGAGCTTATCAAGTTGGCCAAATGTGACCGTGATGGCGGCATT 180
DB      121 CAAAGGGGCTCGTGGAGCTTATCAAGTTGGCCAAATGTGACCGTGATGGCGGCATT 180

QY      181 GGCTTGGGCTTCTGACCTGAGTTTCGAGACACAGCTGGAGAGTGTGGCTTCAAC 240
DB      181 GGCTTGGGCTTCTGACCTGAGTTTCGAGACACAGCTGGAGAGTGTGGCTTCAAC 240

QY      241 CTCTTCAATGCTGGGCTTGGTGTGACATGGGCAATCTGCTGGACGCTTCTTAAGCAG 300
DB      241 CTCTTCAATGCTGGGCTTGGTGTGACATGGGCAATCTGCTGGACGCTTCTTAAGCAG 300

QY      301 TTCCCTTCTGGGAAAGTGTGATCACACTGTTTCAGTATTCGGCTGGCCACCATGAGTGT 360
DB      301 TTCCCTTCTGGGAAAGTGTGATCACACTGTTTCAGTATTCGGCTGGCCACCATGAGTGT 360

QY      361 TTGCTGGTGTGATCTGAGTGTGATGCTGTTGGGGAAGTCACTTGGCGCAGTGTG 420
DB      361 TTGCTGGTGTGATCTGAGTGTGATGCTGTTGGGGAAGTCACTTGGCGCAGTGTG 420

QY      421 GTGATGTGTGCTGGTGGAGGTGACAGCTTTAGGCAACCTGAGAGTGTGATCATGTAATATC 480
DB      421 GTGATGTGTGCTGGTGGAGGTGACAGCTTTAGGCAACCTGAGAGTGTGATCATGTAATATC 480

QY      481 TTCAACACAGACTACACATGATGACATCTAAGTGTGCGAGCTATTTTGGG 540
DB      481 TTCAACACAGACTACACATGATGACATCTAAGTGTGCGAGCTATTTTGGG 540

QY      541 CTGCTGTGGCTGTGGTCTGCAAAAGCTTCTTACCCGAGGGAACGAGATTAAGATCAG 600
DB      541 CTGCTGTGGCTGTGGTCTGCAAAAGCTTCTTACCCGAGGGAACGAGATTAAGATCAG 600

QY      601 ACAGCAACGATACCAAGTTTGTCTGCCATGTGGGGCCCTCTTCTTGTGGATTTCTGG 660
DB      601 ACAGCAACGATACCAAGTTTGTCTGCCATGTGGGGCCCTCTTCTTGTGGATTTCTGG 660

QY      661 CCAAGTTTCAACTGTGCTGCTGAGAAGTGCATGGAAGGAAGAAATGCCGTTCAC 720
DB      661 CCAAGTTTCAACTGTGCTGCTGAGAAGTGCATGGAAGGAAGAAATGCCGTTCAC 720

QY      721 ACCTACTATGTGTAGCAGTCAAGGTGTGACAGCAATCTCAGGGTCACTCTTGGCTCAC 780
DB      721 ACCTACTATGTGTAGCAGTCAAGGTGTGACAGCAATCTCAGGGTCACTCTTGGCTCAC 780

QY      781 CCCCAGGGAAGATCAGCAGACTTATGTGCAAGTGGGTGTGGCAGGAGGGGTGCT 840
DB      781 CCCCAGGGAAGATCAGCAGACTTATGTGCAAGTGGGTGTGGCAGGAGGGGTGCT 840

```

```

QY      841 GTGGGTACCTGATGTACCTGATCCCTTCTCCGTGGCTGCATGTGTGCTTGTG 900
DB      841 GTGGGTACCTGATGTGTACCTGATCCCTTCTCCGTGGCTGCATGTGTGCTTGTG 900

QY      901 GCTGGGTGATCTTCTCGTGGGGAGACCAAGTACTCGCGGGGTGTGTAACGAGTCTG 960
DB      901 GCTGGGTGATCTTCTCGTGGGGAGACCAAGTACTCGCGGGGTGTGTAACGAGTCTG 960

QY      961 GGGATCCCAAGCTGCATATGAGGCTACAACTTCAGCTTGGGGTCTGCTGGAGAG 1020
DB      961 GGGATCCCAAGCTGCATATGAGGCTACAACTTCAGCTTGGGGTCTGCTGGAGAG 1020

QY      1021 ATCATCTACATTTGTGCTGTGCTGTGATACCGTCCGAGCCGCAATGGCATGTTGGC 1080
DB      1021 ATCATCTACATTTGTGCTGTGCTGTGATACCGTCCGAGCCGCAATGGCATGTTGGC 1080

QY      1081 TTCCAGGTCTCTCTCAGCATTTGGGAACTCAGCTTGCCCATCGTATAGCTCTCAGCTCT 1140
DB      1081 TTCCAGGTCTCTCTCAGCATTTGGGAACTCAGCTTGCCCATCGTATAGCTCTCAGCTCT 1140

QY      1141 GGTCCTCGACAGGTTGCTCTTAATCTTAATAATAGAAAGACCTCATAGAGCTTAA 1200
DB      1141 GGTCCTCGACAGGTTGCTCTTAATCTTAATAATAGAAAGACCTCATAGAGCTTAA 1200

QY      1201 TATTTGATGACCAAGTTTCTGGAATTCCTCATTTGGCTGTGGATTTTAA 1254
DB      1201 TATTTGATGACCAAGTTTCTGGAATTCCTCATTTGGCTGTGGATTTTAA 1254

```

```

RESULT 3
HUMRHDAANTI
LOCUS      HUMRHDAANTI 4354 bp mRNA linear PRI 24-AUG-1993
DEFINITION Human Rhd blood group antigen mRNA, complete cds.
ACCESSION L08429.1 GI:337390
VERSION L08429.1 GI:337390
KEYWORDS Rh blood group; Rhd blood group; antigen; blood group antigen.
SOURCE Homo sapiens (library: HL058b (from Clontech)) Bone marrow cDNA to mRNA.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1. (bases 1 to 1354)
AUTHORS Arce,M.A., Thompson,E.S., Wagner,S., Coyne,K.E., Ferdman,B.A. and
Lublin,D.M.
TITLE Molecular cloning of Rhd cDNA derived from a gene present in
JOURNAL Blood 82 (2), 651-655 (1993)
MEDLINE 93320449
PUBMED 8329718
FEATURES
source

```

```

CDS
1..1254
/standard_name="Rhd"
/codon_start=1
/evidence=experimental
/product="Rhd blood group antigen"
/protein_id="AA02679.1"
/db_xref="GI:337391"
/translation="MSSKYPRSYRCLPLMALTLLEAALLLPFYTHYASLEDOGL
VASVYQDPLTWAAIGLGLFSSPRRHSMSVAFLPMLALGVMAIILIDGLSOP
SGRVVITLPSIRLATMSALSVLSVDAVIGKVALQLVVWVLEVTALGNLRVINSI
FNTDYHNMNMHIYVFAVFGLSVAMCLPEFLPEGTBDKQATATPBLISANLGLFLMM
FMSFNSALIRSPDKRKNAVFNTRYAVAVSVVAISGSSLAHQGKISKYVSAVLA
GVAVGTSCHLIPSPMLAMVLGVLGAGISVGAKYLPGCCNRVLGI.PHSIMGNFSL
LGLLGEIYIVLIVLDTVGAGNMIGFOVLSIGELSLAIVIALTSGLLTGLLNLKI
WKAPRAKYPDDQVFWKPRHLAVGF"
BASE COUNT      278 a      346 c      369 g      361 t
ORIGIN

```

Query Match	Similarity	99.9%	Score	1252.4	DB	9	Length	1354
Best Local	Similarity	99.9%	Pred.	No. 0				
Matches	1253	Conservative	0	Mismatches	1	Indels	0	Gaps
QY	1	ATGAGCTCTAAGTACCCGCGGTCTGTCCGCGCCTGCTGCCCCCTTGAGCCCTTAACACTG	60					
Db	1	ATGAGCTCTAAGTACCCGCGGTCTGTCCGCGCCTGCTGCCCCCTTGAGCCCTTAACACTG	60					
QY	61	GAGGAGCTCTCATTTCTCTCTTTTAATTTTTTACCCTATGACGCTTCTTTAGAGAT	120					
Db	61	GAGGAGCTCTCATTTCTCTCTTTTAATTTTTTACCCTATGACGCTTCTTTAGAGAT	120					
QY	121	CAAAAAGGGGCTGTGGCATCTATCAAGTTGGCAAGATCTGACCGTATGGGGGCAAT	180					
Db	121	CAAAAAGGGGCTGTGGCATCTATCAAGTTGGCAAGATCTGACCGTATGGGGGCAAT	180					
QY	181	GCGTTGGGCTTCTCACTCGAGTTTCCGAGACACAGCTGGAGCAGTATGGCTTCAAC	240					
Db	181	GCGTTGGGCTTCTCACTCGAGTTTCCGAGACACAGCTGGAGCAGTATGGGCTTCAAC	240					
QY	241	CTCTTCATGCTGGCGCTTGGTGTGCACTGGGCAATCTCTGTGAACGGCTTCTGAGCCAG	300					
Db	241	CTCTTCATGCTGGCGCTTGGTGTGCACTGGGCAATCTCTGTGACGGCTTCTGAGCCAG	300					
QY	301	TTCCCTTCGGGGAAGGTGTCATCACTCGTTCAAGTATCCGCTGGCCCACTATGATGCT	360					
Db	301	TTCCCTTCGGGGAAGGTGTCATCACTCGTTCAAGTATCCGCTGGCCCACTATGATGCT	360					
QY	361	TTGTGCGTCTCATCTCACTGATGCTGTCTTGGGGAGTCAACTTGGCGAGTTGGTG	420					
Db	361	TTGTGCGTCTCATCTCACTGATGCTGTCTTGGGGAGTCAACTTGGCGAGTTGGTG	420					
QY	421	GTGATGTGCTGTGTGAGGTGACAGCTTTAGGCACTGAGATGTCTATCAATAATATC	480					
Db	421	GTGATGTGCTGTGTGAGGTGACAGCTTTAGGCACTGAGATGTCTATCAATAATATC	480					
QY	481	TTTCAACACAGATACCAACATGAAACATGAGCACTTACGTTCCGACCACTATTTTGGG	540					
Db	481	TTTCAACACAGATACCAACATGAAACATGAGCACTTACGTTCCGACCACTATTTTGGG	540					
QY	541	CTGTCTGTGGCTGTGGTCTGTGCCAAAGCTCTTACCCGAGGAGCAAGAGATTAAGATCAG	600					
Db	541	CTGTCTGTGGCTGTGGTCTGTGCCAAAGCTCTTACCCGAGGAGCAAGAGATTAAGATCAG	600					
QY	601	ACAGCAACGATACCCAGATTGTCTGCCATGTGGGCGCCCTCTTCTTGAGATGTTCTGG	660					
Db	601	ACAGCAACGATACCCAGATTGTCTGCCATGTGGGCGCCCTCTTCTTGAGATGTTCTGG	660					
QY	661	CCAAGTTTCACTTGTCTTCTGTGAGAAATCCAAATGAAAGGAAGATCCCGTGTCAAC	720					
Db	661	CCAAGTTTCACTTGTCTTCTGTGAGAAATCCAAATGAAAGGAAGATCCCGTGTCAAC	720					
QY	721	ACCTACTATGTGTAGCACTCAGCGGTGTGAGCAGCCATCTCAGAGGTCACTCTTGCTCAC	780					
Db	721	ACCTACTATGTGTAGCACTCAGCGGTGTGAGCAGCCATCTCAGAGGTCACTCTTGCTCAC	780					
QY	781	CCCCAAGGAGATCAGCAAGACTTAATGTGCAAGTGCAGTGCAGTGTGGCAGAGGCGTGGCT	840					
Db	781	CCCCAAGGAGATCAGCAAGACTTAATGTGCAAGTGCAGTGCAGTGTGGCAGAGGCGTGGCT	840					
QY	841	GTGGGTAACCTCGTGTCACTGATCCCTTCTCGTGGCTTGGCAATGGTGTGGGTCTTGTG	900					
Db	841	GTGGGTAACCTCGTGTCACTGATCCCTTCTCGTGGCTTGGCAATGGTGTGGGTCTTGTG	900					
QY	901	GCTGGGCTATCTCGTGGGGGAGCCAGATCACTTCCGGGGGTGTGTAACGAGATGCTG	960					
Db	901	GCTGGGCTATCTCGTGGGGGAGCCAGATCACTTCCGGGGGTGTGTAACGAGATGCTG	960					
QY	961	GGGATTTCCACAGCTTCATCATGTGGCTCAACTTCAAGTCTGTGGGTCTGTGGAGAG	1020					
Db	961	GGGATTTCCACAGCTTCATCATGTGGCTCAACTTCAAGTCTGTGGGTCTGTGGAGAG	1020					
QY	1021	ATCATCTACATTGTGCTGTGCTTGTATACCGTGGAGCCGCGCATGATGTGCT	1080					

Db	Accession	Source	Organism	Reference	Authors	Title	Journal	Medline	PubMed	Comment	Features	CDS
Db	1021	ATCATCTACATTTGCTTCTGTCTGTATACCGTCCGAGCGCGCAATGCATATGGC										
Db	1081	TTCCAGTCTCTCCACGATTTGGGAACTACGCTTGGCCATCTGTATAGCTCTCAAGTCT										
Db	1081	TTCCAGTCTCTCCACGATTTGGGAACTACGCTTGGCCATCTGTATAGCTCTCAAGTCT										
Db	1141	GCTCTCCGACAGGTTTGGCTCTAAATCTTAAATATAGGAAGAAGCCATGAGAGCTAAA										
Db	1141	GCTCTCCGACAGGTTTGGCTCTAAATCTTAAATATAGGAAGAAGCCATGAGAGCTAAA										
Db	1201	TATTTTGTATGACCAAGTTTCTGGAAGTTTCCATATTTGGCTGTGATTTTAA										
Db	1201	TATTTTGTATGACCAAGTTTCTGGAAGTTTCCATATTTGGCTGTGATTTTAA										
RESULT 4												
LOCUS	HSRH11											
DEFINITION	H.sapiens mRNA for rheus polypeptide (Rh11).											
ACCESSION	X63094.1											
VERSION	GI:36027											
KEYWORDS	isoform; red cell membrane protein; Rh blood group; Rh polypeptide; Rheus polypeptide.											
SOURCE	Homo sapiens.											
ORGANISM	Homo sapiens.											
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.											
AUTHORS	1 (bases 1 to 1545)											
TITLE	Colin, Y.											
REFERENCE	Submitted (04-NOV-1991) Y. Colin, INSEEM, U76, INTS, 6, Rue											
AUTHORS	A. Cabanel, 75015 Paris, FRANCE											
JOURNAL	2 (bases 1 to 1545)											
REFERENCE	Le Van Kim, C., Cherif-Zahar, B., Raynal, V., Mourou, I., Lopez, M.,											
AUTHORS	Carton, J.P. and Colin, Y.											
TITLE	Multiple Rh messenger RNA isoforms are produced by alternative											
JOURNAL	splicing											
MEDLINE	Blood 80 (4), 1074-1078 (1992)											
PUBMED	92360855											
COMMENT	1379850											
FEATURES	3 (bases 1 to 1545)											
REFERENCE	Le van Kim, C., Mourou, I., Cherif-Zahar, B., Raynal, V., Cherrier, C.,											
AUTHORS	Carton, J.P. and Colin, Y.											
TITLE	Molecular cloning and primary structure of the human blood group											
JOURNAL	RhD polypeptide											
MEDLINE	Proc. Natl. Acad. Sci. U.S.A. 89 (22), 10925-10929 (1992)											
PUBMED	93066356											
COMMENT	1438298											
FEATURES	See also X63094.98 & M34015.											
REFERENCE	Location/Qualifiers											
FEATURES	1. 1545											
REFERENCE	/organism="Homo sapiens"											
FEATURES	/db_xref="taxon:9606"											
REFERENCE	/chromosome="1"											
FEATURES	/map="1p35"											
REFERENCE	/clone="Rh11"											



GVAVGTSCHLIPSPMLAMVLGLVAGLISVGAKYLPGCCNRVLGIPHSIMGNPSL  
 LGALGEIYIVLVDITVGAGNGMIGQVLSIGELSLAIVIALTSLGLTLNLTKI  
 WKAPHEKPYDDQVFMKFPHLAVGF"  
 BASE COUNT 657 a 680 c 652 g 801 t  
 ORIGIN

Query Match 99.9%; Score 1252.4; DB 9; Length 2790;  
 Best Local Similarity 99.9%; Pred. No. 0;  
 Matches 1253; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGAGCTTAAGTACCCGGGCTGTCCGGGCGTGGCCGCCCTGTGGCCCTTAACATG 60  
 DB 17 ATGAGCTTAAGTACCCGGGCTGTCCGGGCGTGGCCGCCCTGTGGCCCTTAACATG 76  
 QY 61 GAAGAGCTCTCATCT 120  
 DB 77 GAAGAGCTCTCATCT 136  
 QY 121 CAAAAGGGGCTGTGGCATCTCTATCAAGTTGGCCAAATCTTGACCGGTATGGCGCCATT 180  
 DB 137 CAAAAGGGGCTGTGGCATCTCTATCAAGTTGGCCAAATCTTGACCGGTATGGCGCCATT 196  
 QY 181 GCGTTGGGCTTCTCACTCCAGTTCCGGAGACACACTGGAGAGATGTGGCCCTTCAAC 240  
 DB 197 GCGTTGGGCTTCTCACTCCAGTTCCGGAGACACACTGGAGAGATGTGGCCCTTCAAC 256  
 QY 241 CTCTTCATGCTGGGCTGT 300  
 DB 257 CTCTTCATGCTGGGCTGT 316  
 QY 301 TTCCCTTGTGGAGAGT 360  
 DB 317 TTCCCTTGTGGAGAGT 376  
 QY 361 TTTCGGGCTGT 420  
 DB 377 TTTCGGGCTGT 436  
 QY 421 GTGATGT 480  
 DB 437 GTGATGT 496  
 QY 481 TTCAACACAGACTACCATGAATGATGACATCTACGTGTGTGTGTGTGTGTGTGTGT 540  
 DB 497 TTCAACACAGACTACCATGAATGATGACATCTACGTGTGTGTGTGTGTGTGTGTGT 556  
 QY 541 CTGT 600  
 DB 557 CTGT 616  
 QY 601 ACAGCAACGATACCCAGTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 660  
 DB 617 ACAGCAACGATACCCAGTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 676  
 QY 661 CCAAGTTTCACTGTCTGT 720  
 DB 677 CCAAGTTTCACTGTCTGT 736  
 QY 721 ACCATCTATGT 780  
 DB 737 ACCATCTATGT 796  
 QY 781 CCCCAAGGAAGATCAGCAAGACTTATGTGACAGTGTGTGTGTGTGTGTGTGTGTGTGT 840  
 DB 797 CCCCAAGGAAGATCAGCAAGACTTATGTGACAGTGTGTGTGTGTGTGTGTGTGTGTGT 856  
 QY 841 GTGGTACCTGT 900  
 DB 857 GTGGTACCTGT 916  
 QY 901 GCTGGGCTGATTCCTGTGGGGGAGCCAAATCTTGGCGGGGTGTGTGTGTGTGTGTGT 960  
 DB 917 GCTGGGCTGATTCCTGTGGGGGAGCCAAATCTTGGCGGGGTGTGTGTGTGTGTGTGT 976

QY 961 GGAATCCCAACAGCTTCATCATGGGCTACAACTTACGCTTGTGGTCTGTGGAGAG 1020  
 DB 977 GGAATCCCAACAGCTTCATCATGGGCTACAACTTACGCTTGTGGTCTGTGGAGAG 1036  
 QY 1021 ATCATCTACATTTGT 1080  
 DB 1037 ATCATCTACATTTGT 1096  
 QY 1081 TTCAGGCTCTCTCAGCATTTGGGGAATCAGTTGGCCATCGATAGCTTCACGCTCT 1140  
 DB 1097 TTCAGGCTCTCTCAGCATTTGGGGAATCAGTTGGCCATCGATAGCTTCACGCTCT 1156  
 QY 1141 GGTCTCTGACAGGTTTGTCTCTAAATCTTAAATATGAAAGCCTCATGAGCTTAAA 1200  
 DB 1157 GGTCTCTGACAGGTTTGTCTCTAAATCTTAAATATGAAAGCCTCATGAGCTTAAA 1216  
 QY 1201 TATTTTGTATGACCAAGTTTCTGGAAGTTTCTCATTTGGCTGTGGATTTTAA 1254  
 DB 1217 TATTTTGTATGACCAAGTTTCTGGAAGTTTCTCATTTGGCTGTGGATTTTAA 1270

RESULT 6  
 AB018966 1254 bp mRNA linear (PRI 03-JUN-1999)  
 LOCUS  
 DEFINITION  
 Homo sapiens RhDva (FK) mRNA for Rh blood group D antigen (Rhd), complete cds.  
 ACCESSION  
 AB018966 GI:5360242  
 VERSION  
 RhDva (FK); Rh blood group D antigen (Rhd).  
 KEYWORDS  
 Homo sapiens (Isolate:FK) cDNA to mRNA.  
 SOURCE  
 ORGANISM  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE  
 1 (sites)  
 Hyodo, H., Ishikawa, Y., Kashiwase, K., Ogawa, A., Watanabe, Y.,  
 Taneyama, H., Toyoda, C., Uchikawa, M., Akaza, T. and Fujii, T.  
 Polymorphisms of RhDva in Japanese  
 Unpublished  
 2 (bases 1 to 1254)  
 Uchikawa, M., Hyodo, H. and Ishikawa, Y.  
 Direct Submission  
 Submitted (20-OCT-1998) Hironobu Hyodo, Japanese Red Cross Central  
 Blood Center, Research; 4-1-31, Hiroo, Shibuya, Tokyo 150-0012,  
 Japan (E-mail:hyodo@hia.abc.jrc.or.jp, Tel: +81-3-5485-6009,  
 Fax: +81-3-3406-7892)

## FEATURES

source  
 1..1254  
 Location/Qualifiers  
 /organism="Homo sapiens"  
 /isolate="FK"  
 /db\_xref="taxon:9606"  
 1..1254  
 /gene="RhDva (FK)"  
 1..1254  
 /gene="RhDva (FK)"  
 /note="Rhd variant RhDva (FK)"  
 /codon\_start=1  
 /product="Rh blood group D antigen (Rhd)"  
 /protein\_id="BA81899.1"  
 /db\_xref="GI:5360243"  
 /translation="MSSKYPKPRVRCPLMALTLBAALILFYPFTHYDASLEDEKCL  
 VASYQVODLTVAAGLGLFTSSPRHSVSAFNLFMALGVQVAILLDGLSQP  
 SGKVVLTLEFIRLATWSALSVLISVDVAGKVMIAQIVVAVLEVTAGMLRWISNI  
 FNDYHNMMHIVPAFAPGLSVAMCKPKPLPGTEBDKDTATLPSISAMIGALFLVM  
 FMSVNSALLRSPIORKNVFNTRYVAVSVVAVRAIGSSLAHPGKISKTYHSAVLA  
 GYRAVGTSCHLIPSPMLAMVLGLVAGLISVGAKYLPGCCNRVLGIPHSIMGNPSL  
 LGALGEIYIVLVDITVGAGNGMIGQVLSIGELSLAIVIALTSLGLTLNLTKI  
 WKAPHEKPYDDQVFMKFPHLAVGF"

BASE COUNT 241 a 323 c 351 g 339 t  
 ORIGIN  
 Query Match 99.7%; Score 1250.8; DB 9; Length 1254;  
 Best Local Similarity 99.8%; Pred. No. 0;



Matches 1252; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

QY 1 ATGAGCTCTAAGTACCCGCGGTCTGTCCGGCTGCTGCCCTCTGTGGCCCTTAACACTG 60
Db 1 ATGAGCTCTAAGTACCCGCGGTCTGTCCGGCTGCTGCCCTCTGTGGCCCTTAACACTG 60
QY 61 GAAGACCTCTCATTTCTCTCTTTATTTTAAACCACTATGACGCTTCTTGAAGAT 120
Db 61 GAAGACCTCTCATTTCTCTCTTTATTTTAAACCACTATGACGCTTCTTGAAGAT 120
QY 121 CAAAGGGGCTCGTGGAGCTCATCAAGTTGGGCAAGTCCGAGTGGGGGCAATT 180
Db 121 CAAAGGGGCTCGTGGAGCTCATCAAGTTGGGCAAGTCCGAGTGGGGGCAATT 180
QY 181 GCGTTGGGCTTCTCACTCGAGTTTCCGAGACACAGCTGGACAGTGGCCCTTCAAC 240
Db 181 GCGTTGGGCTTCTCACTCGAGTTTCCGAGACACAGCTGGACAGTGGCCCTTCAAC 240
QY 241 CTCTTCAATGCTGGCGCTTGGTGTGCAAGTGGGCAATCCTGCTGACCGCTTCTGAGCCAG 300
Db 241 CTCTTCAATGCTGGCGCTTGGTGTGCAAGTGGGCAATCCTGCTGAGCGCTTCTGAGCCAG 300
QY 301 TTCCCTCTGGGAGGAGTGCATGACACTGTTCAGTATTTGGCTGGCCACCATAGAGTCT 360
Db 301 TTCCCTCTGGGAGGAGTGCATGACACTGTTCAGTATTTGGCTGGCCACCATAGAGTCT 360
QY 361 TTGCGGCTGCTGATCTAGTGGATGTGTCTTGGGGAAGTCACTTGGCGCAGTTGGGTG 420
Db 361 TTGCGGCTGCTGATCTAGTGGATGTGTCTTGGGGAAGTCACTTGGCGCAGTTGGGTG 420
QY 421 GTGATGCTGTGGTGGAGTGAAGCTTTTAAAGCACTTGAAGTGTATCATGATATATC 480
Db 421 GTGATGCTGTGGTGGAGTGAAGCTTTTAAAGCACTTGAAGTGTATCATGATATATC 480
QY 481 TTGAACACACATCAACCAATGAATGATGACATCACTGCTTGGAGCCTATTTTGGG 540
Db 481 TTGAACACACATCAACCAATGAATGATGACATCACTGCTTGGAGCCTATTTTGGG 540
QY 541 CTGCTGTGGCTGTGGTCTGTCGCAAGCTTCAACCGAGGGAAGGAGATTAAGATCAG 600
Db 541 CTGCTGTGGCTGTGGTCTGTCGCAAGCTTCAACCGAGGGAAGGAGATTAAGATCAG 600
QY 601 ACAGCAACGATACCCAGTTGTCTGCAATGCTGGGCGCCCTTCTTGTGTGATGTTCTGG 660
Db 601 ACAGCAACGATACCCAGTTGTCTGCAATGCTGGGCGCCCTTCTTGTGTGATGTTCTGG 660
QY 661 CCAAGTTCAACTCTGCTCTGCTGAGAACTCAATCGAAAGAAATGCGGTGTTCAAC 720
Db 661 CCAAGTTCAACTCTGCTCTGCTGAGAACTCAATCGAAAGAAATGCGGTGTTCAAC 720
QY 721 ACTACTATGCTGTAGAGTGAAGGTGTGACAGCATCTCAGGGTCAATCTTGGGCTCAC 780
Db 721 ACTACTATGCTGTAGAGTGAAGGTGTGACAGCATCTCAGGGTCAATCTTGGGCTCAC 780
QY 781 CCCCAGGGAAGATCAGCAAGCTTATGTGCAAGTGGGTGTTGGCAGAGGCGTGGCT 840
Db 781 CCCCAGGGAAGATCAGCAAGCTTATGTGCAAGTGGGTGTTGGCAGAGGCGTGGCT 840
QY 841 GTGGGTACCTGTGTCACTGATCCCTTCTCGGTGGCTTGGCAATGCTGGGTCTTGTG 900
Db 841 GTGGGTACCTGTGTCACTGATCCCTTCTCGGTGGCTTGGCAATGCTGGGTCTTGTG 900
QY 901 GCTGGGCTGATCTCGTGGGGGAGCCAAAGTACCTGGGGGTGTGTAACGAGTGGTG 960
Db 901 GCTGGGCTGATCTCGTGGGGGAGCCAAAGTACCTGGGGGTGTGTAACGAGTGGTG 960
QY 961 GGGATTCCTCAACAGCTTCATGAGGCTTCAAACTTCAAGCTTGTGGGTCTTGTGAGAG 1020
Db 961 GGGATTCCTCAACAGCTTCATGAGGCTTCAAACTTCAAGCTTGTGGGTCTTGTGAGAG 1020
QY 1021 ATCATCTACATTTGCTGTGTGCTTGTATACCGTGGAGCGGCAATGAGATGATTTGGC 1080
Db 1021 ATCATCTACATTTGCTGTGTGCTTGTATACCGTGGAGCGGCAATGAGATGATTTGGC 1080

```

```

QY 1081 TTCCAGGTCTCTCAGACATTTGGGAACTCAGCTTGGCCATCGTATAGCTTCCAGCTCT 1140
Db 1081 TTCCAGGTCTCTCAGACATTTGGGAACTCAGCTTGGCCATCGTATAGCTTCCAGCTCT 1140
QY 1141 GGTCTCTGACAGGTTGCTTCTTAAATCTTAAATATGAAAGCACTCATGAGGCTAAA 1200
Db 1141 GGTCTCTGACAGGTTGCTTCTTAAATCTTAAATATGAAAGCACTCATGAGGCTAAA 1200
QY 1201 TATTTATGACCAAGTTTCTGGAAGTTTCTCATTTGGCTGTGATTTTAA 1254
Db 1201 TATTTATGACCAAGTTTCTGGAAGTTTCTCATTTGGCTGTGATTTTAA 1254

```

RESULT 7  
AB018967  
LOCUS  
DEFINITION  
Homo sapiens RHDVa (TO) mRNA for Rh blood group D antigen (RHD), complete cds.

ACCESSION  
AB018967  
VERSION  
AB018967.1 GI:5360244  
KEYWORDS  
RHDVa (TO); Rh blood group D antigen (RHD).  
SOURCE  
Homo sapiens (isolate:TO) cDNA to mRNA.  
ORGANISM  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

REFERENCE  
AUTHORS  
Hyodo, H., Ishikawa, Y., Kashiwase, K., Ogawa, A., Watanabe, Y., Tsuneyama, H., Toyoda, C., Uchikawa, M., Akaza, T. and Fujii, T.  
TITLE  
JOURNAL  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
Submitted (20-OCT-1999) Hironobu Hyodo, Japanese Red Cross Central Blood Center, Research; 4-1-31, Hiroo, Shibuya, Tokyo 150-0012, Japan (E-mail:hyodo@hla.obc.jrc.or.jp, Tel:+81-3-5485-6009, Fax:+81-3-3406-7892)

FEATURES  
source  
1..1254  
/organism="Homo sapiens"  
/isolate="TO"  
/db\_xref="taxon:9606"  
1..1254  
/gene="RHDVa (TO)"  
1..1254  
/gene="RHDVa (TO)"  
/note="Rhd variant RHDVa (TO)"  
/codon\_start=1  
/product="Rh blood group D antigen (RHD)"  
/protein\_id="BA081900.1"  
/db\_xref="GI:5360245"  
/translation="MSSKTPRSVRLCPMLATLLEALILFFPHHYDASLEDQKGLVASVYQGDUTVMAALIGLFTSFRHSMSSVAFLFMALGVQNALIDGLSQFP SGRKVIYTLPSIRLATMSALSVLSVDVAGKVLADLVVAVLEVALIRVVISNI FNDYVIMMMHIIYVPAFAPGLSVAKLPEPLPGTDDKQDTATIPSLMGLPLMM FWPSPNSALRSPIDRKNAMENTYVAVSVYTAIIGSSLAHPGKISKTYSAVLA GSAVSTSCHLIPSPMLAVLGLISVGAKEYIPGCCNRLVGIPIHSHIMYNSVL LGLVGLIYIVLVDPTVAGNEMICFPQVLLISIGLSLALVALTGLTGLILNKKI WKAPHPAKYFDVDVFWKPHLAVGF"

BASE COUNT  
242 a 349 g 340 c  
ORIGIN  
Query Match 99.7%; Score 1250.8; DB 9; Length 1254;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 1252; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

QY 1 ATGAGCTCTAAGTACCCGCGGTCTGTCCGGCTGCTGCCCTCTGTGGCCCTTAACACTG 60
Db 1 ATGAGCTCTAAGTACCCGCGGTCTGTCCGGCTGCTGCCCTCTGTGGCCCTTAACACTG 60
QY 61 GAAGACCTCTCATTTCTCTCTTTATTTTAAACCACTATGACGCTTCTTGAAGAT 120
Db 61 GAAGACCTCTCATTTCTCTCTTTATTTTAAACCACTATGACGCTTCTTGAAGAT 120

```

Db	61	GAAGACGCTCATTCTCTCTTCAATTTTTCACCACTATGACGCTTCTTAAGAGAT	120
Qy	121	CAAAAGGGGCTCTGTCGACTCTCTATCAAGTTGGCCAAAGATCTGACCTGATGGGGCATT	180
Db	121	CAAAAGGGGCTCTGTCGACTCTCTATCAAGTTGGCCAAAGATCTGACCTGATGGGGCATT	180
Qy	181	GGCTTGGGCTTCTCACCCTCGAGTTTCCGGAGACACAGCTGGAGCAGTGTGGCTTCAAC	240
Db	181	GGCTTGGGCTTCTCACCCTCGAGTTTCCGGAGACACAGCTGGAGCAGTGTGGCTTCAAC	240
Qy	241	CTCTTCATGCTGCGGCTTGTGTGACAGTGGGCAATCTCTGTGACGGCTTCTGAGCCAG	300
Db	241	CTCTTCATGCTGCGGCTTGTGTGACAGTGGGCAATCTCTGTGACGGCTTCTGAGCCAG	300
Qy	301	TTCCCTTCTGGGAGAGTGTGATCACTGTTCAATTTGGGCTGGCCACCATGATGCT	360
Db	301	TTCCCTTCTGGGAGAGTGTGATCACTGTTCAATTTGGGCTGGCCACCATGAGGCT	360
Qy	361	TTGTGGGTCGATCTCAGTGGATGCTGTGGGGAAAGGCACTTGGGCGAGTTGGTG	420
Db	361	TTGTGGGTCGATCTCAGTGGATGCTGTGGGGAAAGGCACTTGGGCGAGTTGGTG	420
Qy	421	GTGATGTGCTGCTGGTGGAGGTGACAGCTTTAGGCAACCTGAGAGTGCATCAAGTAATC	480
Db	421	GTGATGTGCTGCTGGTGGAGGTGACAGCTTTAGGCAACCTGAGAGTGCATCAAGTAATC	480
Qy	481	TTCAACACAGACTACACATGAAACATGACATCTACGTGTTCGAGCCTATTTTGGG	540
Db	481	TTCAACACAGACTACCAATGAAACATGATGACATCTACGTGTTCGAGCCTATTTTGGG	540
Qy	541	CTGTCTGTGGCTGGTGGCTGCGCAAGGCTTTACCCAGGGAAACGAGAGATAAAGATCAG	600
Db	541	CTGTCTGTGGCTGGTGGCTGCGCAAGGCTTTACCCAGGGAAACGAGAGATAAAGATCAG	600
Qy	601	ACAGCAAGATACCCAGTTTGTCTGCAATGCTGGAGCGCCCTTCTTGTAGATGTTCTGG	660
Db	601	ACAGCAAGATACCCAGTTTGTCTGCAATGCTGGAGCGCCCTTCTTGTAGATGTTCTGG	660
Qy	661	CCAAAGTTTCAACTGCTCTGCTGAGAAAGTCCAAATCGAAAGAAATGCGGTGTTCAAC	720
Db	661	CCAAAGTTTCAACTGCTCTGCTGAGAAAGTCCAAATCGAAAGAAATGCGATGTTCAAC	720
Qy	721	ACCTACTATGCTGTAGCAGTACAGCTGGTGAACGCATCTGAGGGTCACTTGGCTCAC	780
Db	721	ACCTACTATGCTGTAGCAGTACAGCTGGTGAACGCATCTGAGGGTCACTTGGCTCAC	780
Qy	781	CCCAAGGGAAGATGAGAAAGCTTATGTGACAGTGGCGGTGGTGGAGGAGCGTGGCT	840
Db	781	CCCAAGGGAAGATGAGAAAGCTTATGTGACAGTGGCGGTGGTGGAGGAGCGTGGCT	840
Qy	841	GTGGGTACCTCGTGCACCTGATCCCTCTCCGTGGCTGTGCATGCTGTGGTCTTGTG	900
Db	841	GTGGGTACCTCGTGCACCTGATCCCTCTCCGTGGCTGTGCATGCTGTGGTCTTGTG	900
Qy	901	GCTGGGCTGATCTCCGTGGGGGAGCCAAAGTACCTGCGGGGTGTGTAACCCGAGTGC	960
Db	901	GCTGGGCTGATCTCCGTGGGGGAGCCAAAGTACCTGCGGGGTGTGTAACCCGAGTGC	960
Qy	961	GGGATTTCCCAAGCTCATATATGGGCTCAAACTTCAAGCTTGTGGGTCTGTGGAGAG	1020
Db	961	GGGATTTCCCAAGCTCATATATGGGCTCAAACTTCAAGCTTGTGGGTCTGTGGAGAG	1020
Qy	1021	ATCATCTACATGTCTCTGTGGTCTTATATACCGTGGAGCGGCAATGGCATATTTGGC	1080
Db	1021	ATCATCTACATGTCTCTGTGGTCTTATATACCGTGGAGCGGCAATGGCATATTTGGC	1080
Qy	1081	TTCCAGGTCCTCTCAGCATTTGGGGAAGCTTGGCCATCTGTATAGCTTCAAGCTT	1144
Db	1081	TTCCAGGTCCTCTCAGCATTTGGGGAAGCTTGGCCATCTGTATAGCTTCAAGCTT	1144
Qy	1141	GCTTCTCTGACAGGTTTGTCTCTTAATTTTAAATATAGGAAGACCTCATGAGGCTTAA	1200
Db	1141	GCTTCTCTGACAGGTTTGTCTCTTAATTTTAAATATAGGAAGACCTCATGAGGCTTAA	1200

LOCUS	578509	1260 bp	MRNA	linear	PRI 27-SEP-1995
DEFINITION	RHD (human, D--phenotype, erythrocyte, mRNA Partial, 1260 nt).				
ACCESSION	J78509				
VERSION	J78509.1				
KEYWORDS	GI:999309				
SOURCE	Homo sapiens				
ORGANISM	Homo sapiens erythrocyte D-phenotype.				
REFERENCE	Bukacinski, Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 1260)				
AUTHORS	Huang, C.H., Reid, M.E. and Chen, Y.				
TITLE	Identification of a partial internal deletion in the RH locus causing the human erythrocyte D-phenotype				
JOURNAL	Blood 86 (2), 784-790 (1995)				
MEDLINE	95329738				
PUBMED	7606008				
REMARK	Genbank staff at the National Library of Medicine created this entry [NCBI gibbs 16818] from the original journal article. This sequence comes from Fig. 6.				
FEATURES	Location/Qualifiers				
source	1..1260				
gene	/organism="Homo sapiens"				
CDS	/db_xref="taxon:9606"				
	1..1260				
	/gene="RHD"				
	1..1254				
	/gene="Rhd"				
	/note="This sequence comes from Fig. 6"				
	/codon_start=1				
	/protein_id="AAB34852.1"				
	/db_xref="GI:999310"				
	/translation="MSSKVPBSRVSRKPLCALTLLEALILLFPTHYDASLEPOKGLVASIQVQDILTVMAATLGLFISFRPHSSSVAFNLFMLALGQWAILDGLFSGRP SGKAVITLFSIRLATWSALTSVLISVDVLKVNLAQLVWLVLEVTALGNDRVTSIN FNDYNNHMMHITWFAFVFGLSVAMCLPKPLPETEDKDDPATIPSLANLGAFLWMM FWPBFSNALLRSPERKNAFENTYAVAVSVVVAISGSSLAHPQKISKTYVSAVLA GGVAVTSCILIPSPMLAMVGLVAGLISGVAKYLEGCNRYVIGIPSSIMGVFSL LGILGELIYVILVLDIVTGAGNGMGISQVLLSIGELSLAIVIALTSLHLLNLEKI WKAPHAKYFDDOVFKKPEPLAVGF"				
BASE COUNT	245 a 325 c 349 g 341 t				
ORIGIN					
Query Match	99.5%; Score 1247.6; DB 9; Length 1260;				
Best Local Similarity	99.7%; Pred. No. 0;				
Matches 1250; Conservative	0; Mismatches 4; Indels 0; Gaps 0;				
1	ATGAGCTCTAAGTACCCCGGCTCTGTCGGCGGCTGCTGCCCTCTGAGGACCTTAACACTG 60				
1	ATGAGCTCTAAGTACCCCGGCTCTGTCGGCGGCTGCTGCCCTCTGAGGACCTTAACACTG 60				
61	GAGCAGCTCTCATCTCTCTCTTCTATATTTTAAACCACTATAGAGCTTCTCTTAGAGAT 120				
61	GAGCAGCTCTCATCTCTCTCTTCTATATTTTAAACCACTATAGAGCTTCTCTTAGAGAT 120				
121	CAAAAGGGGCTCGTGGCATCTATCAATGTGGCCCAAGATCGACCGTATGGCGGCATT 180				
121	CAAAAGGGGCTCGTGGCATCTATCAATGTGGCCCAAGATCTGACCGTATGGCGGCATT 180				
181	GGCTTGGGCTTCTGACCTGAGATTCCGGAGACACAGCTGAGAGCATGTGGCTTCAAC 240				
181	GGCTTGGGCTTCTGACCTGAGATTCCGGAGACACAGCTGAGAGCATGTGGCTTCAAC 240				
241	CTCTTCATGCTGGCGCTTGCTGTGACAGTGGCAATCCCTGCTGAGACGGCTTCTAGACGAG 300				
241	CTCTTCATGCTGGCGCTTGCTGTGACAGTGGCAATCCCTGCTGAGACGGCTTCTAGACGAG 300				
300	CTCTTCATGCTGGCGCTTGCTGTGACAGTGGCAATCCCTGCTGAGACGGCTTCTAGACGAG 300				

```

QY 301 TTCCCTTCTGGGAAGTGTGATCACTGTTCAGTATTCGGCTGGCCACCAGTGTCT 360
Db 301 TTCCCTTCTGGGAAGTGTGATCACTGTTCAGTATTCGGCTGGCCACCAGTGTCT 360
QY 361 TTGTGCGTGTGATCTCACTAGTGTGTCTTGGGGAAGGTCACTTGGCCAGTGTG 420
Db 361 TTGTGCGTGTGATCTCACTAGTGTGTCTTGGGGAAGGTCACTTGGCCAGTGTG 420
QY 421 GTATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 480
Db 421 GTATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 480
QY 481 TTCAACACAGCTACCAATGATGATGATGATGATGATGATGATGATGATGATGAT 540
Db 481 TTCAACACAGCTACCAATGATGATGATGATGATGATGATGATGATGATGATGAT 540
QY 541 CTGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 600
Db 541 CTGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 600
QY 601 ACAGCAAGATACCAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 660
Db 601 ACAGCAAGATACCAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 660
QY 661 CCAAGTTTCACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 720
Db 661 CCAAGTTTCACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 720
QY 721 ACCATCTATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 780
Db 721 ACCATCTATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 780
QY 781 CCCCAAGGAAGATCAGAGACTTATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 840
Db 781 CCCCAAGGAAGATCAGAGACTTATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 840
QY 841 GTGGGTACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 900
Db 841 GTGGGTACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 900
QY 901 GCTGGGTGATCTCCGTGGGGGAGCCAGTACCTGCGGGGGTGTGTACCGAGTGTG 960
Db 901 GCTGGGTGATCTCCGTGGGGGAGCCAGTACCTGCGGGGGTGTGTACCGAGTGTG 960
QY 961 GGGATGCCCAAGCTCATGATGGGCTACCACTTCAAGCTTGGGGCTTGGAGAG 1020
Db 961 GGGATGCCCAAGCTCATGATGGGCTACCACTTCAAGCTTGGGGCTTGGAGAG 1020
QY 1021 ATCATCTACATTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1080
Db 1021 ATCATCTACATTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1080
QY 1081 TTCCAGTCTCTCTGAGCATTTGGGAACTCACTTGGGCTGTGTGTGTGTGTGTGT 1140
Db 1081 TTCCAGTCTCTCTGAGCATTTGGGAACTCACTTGGGCTGTGTGTGTGTGTGTGT 1140
QY 1141 GGTCTCCGACAGGTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1200
Db 1141 GGTCTCCGACAGGTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1200
QY 1201 TATTTTATGACCAAGTTTCTGGAAGTTTCTCATTTGGCTGTGTGTGTGTAA 1254
Db 1201 TATTTTATGACCAAGTTTCTGGAAGTTTCTCATTTGGCTGTGTGTGTGTAA 1254

```

```

RESULT 9
AB018968 AB018968 1254 bp mRNA linear PRI 03-JUL-1999
LOCUS Homo sapiens RhdVa (TT) mRNA for Rh blood group D antigen (RHD),
DEFINITION complete cds.
ACCESSION AB018968
VERSION AB018968.1 GI:5360246

```

```

KEYWORDS RhdVa (TT) ; Rh blood group D antigen (RHD) .
SOURCE Homo sapiens (isolate:TT) cDNA to mRNA.
ORGANISM Homo sapiens
REFERENCE 1 (sites)
AUTHORS Hyodo,H., Ishikawa,Y., Kashiwase,K., Ogawa,A., Watanabe,Y.,
Tsuneyama,H., Toyoda,C., Uchikawa,M., Akaza,T. and Fujii,T.
TITLE Polymorphisms of RhdVa in Japanese
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1254)
AUTHORS Uchikawa,M., Hyodo,H. and Ishikawa,Y.
TITLE Direct Submision
JOURNAL Submitted (20-OCT-1998) Hironobu Hyodo, Japanese Red Cross Central
Blood Center, Research; 4-1-31, Hirono, Shibuya, Tokyo 150-0012,
Japan (E-mail: hyodo@nha.cbc.jrc.or.jp, Tel: +81-3-5485-6009,
Fax: +81-3-3406-7892)
FEATURES
source
1..1254
/organism="Homo sapiens"
/isolate="TT"
/db_xref="taxon:9606"
1..1254
/gene="RhdVa (TT)"
/notes="Rhd variant RhdVa (TT)"
/product="Rh blood group D antigen (RHD)"
/protein_id="BA01901.1"
/db_xref="GI:5360247"
/translation="MSSKYPKRSVRCLPLMALTLLEAAILLPFFTHYDASLDOKGL
VASVOYGDLTWMAIGLGLTSSFRHSVAVFNLPMALVDMALDGLPSOP
SGKVVTFTSIRLATMSALSVLSVAVIGKVLADLVVAVLVEVVALGLRVLVSI
PNTDYAMNMHIVPAHYGLSLVAVLCLPRLPGTBDKDTATIPSLAMGLPLFM
PMSVNSALRSPIDRKAMFNLYVALVAVLISGLSLHPOGKISKTYHSAVLA
GGVAVGTSCHLIPSPWLAVALGLISVGAHYLPCCNRYLGIPIHSSIMYNSL
LGLGLIIVLVLDVTGAGNMIGFQVLLSTGSLAIVLALVGLTGLLNLKI
MKAPHEAKYFDQVPMKPHPLAVGF"
BASE COUNT 242 a 323 c 349 g 340 t
ORIGIN
Query Match 99.4%; Score 1246; DB 9; Length 1254;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 1249; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
QY 1 ATGAGCTCTAAGTACCCGCGGTGTGTCCGGCTGCTGCCCTCTGGGCCCTTAACATG 60
Db 1 ATGAGCTCTAAGTACCCGCGGTGTGTCCGGCTGCTGCCCTCTGGGCCCTTAACATG 60
QY 61 GAAGCAGCTCATTCCTTCCTCTTATTTTAAACCATGAGCGCTTCTTAGAGAT 120
Db 61 GAAGCAGCTCATTCCTTCCTCTTATTTTAAACCATGAGCGCTTCTTAGAGAT 120
QY 121 CAAAAGGGGCTCGTGGCATCTCAAGTTGGCCCAAGATGACCGTATGCGCGCAT 180
Db 121 CAAAAGGGGCTCGTGGCATCTCAAGTTGGCCCAAGATGACCGTATGCGCGCAT 180
QY 181 GGGTGGGCTTCTCACTGAGTTTCCGAGACACAGCTGAGCAGTGTGGCTTCAAC 240
Db 181 GGGTGGGCTTCTCACTGAGTTTCCGAGACACAGCTGAGCAGTGTGGCTTCAAC 240
QY 241 CTCTTCATGCTGGGCTTGGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 300
Db 241 CTCTTCATGCTGGGCTTGGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 300
QY 301 TTCCCTTCTGGGAAGTGTGATCACTGTTCAGTATTCGGCTGGCCACCAGTGTCT 360
Db 301 TTCCCTTCTGGGAAGTGTGATCACTGTTCAGTATTCGGCTGGCCACCAGTGTCT 360
QY 361 TTGTGCGTGTGATCTCACTAGTGTGTCTTGGGGAAGGTCACTTGGCCAGTGTG 420
Db 361 TTGTGCGTGTGATCTCACTAGTGTGTCTTGGGGAAGGTCACTTGGCCAGTGTG 420

```

[illegible]

Journal MEDLINE PubMed REMARK	Category IIIC phenotype Transfusion 36 (6), 567-574 (1996) 8669091 GenBank staff at the National Library of Medicine created this entry [NCBI gidsb 178403] from the original journal article. This sequence comes from Fig. 4.	FEATURES SOURCE
	1..1254 /organism="Homo sapiens" /db_xref="taxon:9606" /cell_type="reticulocyte" /note="alioanti-D producing D-positive patient"	gene
	1..1254 /gene="RHD"	gene
	1..1254 /gene="RHD" /function="associated with hemolytic disease of the newborn in a D-positive child" /note="This sequence comes from Fig. 4; conceptual translation presented here differs from translation in publication"	gene
	/codon_start=1 /product="D category IIIC antigen" /protein_id="AAB37696.2" /db_xref="GI:9790343"	gene
	/translation="MSKSRPRSVRRCPLPLALITLALILFFFTHYDSLNDOKL VAVSQQDILVMAALIGFLTSSFRHSHSVAFNFMALGVQVALIDGLFSQRP SGKVLITFSIRLATMSAMVLISAGVALIKVLAQVAVLVEVTLIRVVISNI FNDYHNMMHIIYFAAYFGISVAMCKPLPTEDEKDOTATIPSLAMIGALFLMM FMEFNSALRSPERKNAVFNRYAVAVSVITLSSSLAHPOGKISKYHSAVALA GGAVGTSCHLIPSPMLAMVLGLVAGLISVGAKYLDGCCRRVLGTHSSIMGNFSL LGILGEIYIVLVLDIVTGAGNGIMIGQVLLISIGLSLAIVALITSLTGLNLTNTKI WKAEHAKYFPDQVFMFEPHIAVGF"	gene
BASE COUNT	240 a 323 c 352 g 339 t	ORIGIN
Query Match	99.4%; Score 1246; DB 9; Length 1254;	
Best Local Similarity	99.6%; Pred. No. 0;	
Matches 1249; Conservative	0; Mismatches 5; Indels 0; Gaps 0;	
1	ATGAGCTCTAAGTACCGGCGGGCTGTCCGGGGGCTGCTGCCCTCTGGAGCCCTTAACATCTG 60	
1	ATGAGCTCTAAGTACCGGCGGGCTGTCCGGGGGCTGCTGCCCTCTGGAGCCCTTAACATCTG 60	
61	GAAAGAGCTCTCATTTCTCCTCTTCTATTTTTTTTAAACCACTATGAGCGTTCTTAGAGAT 120	
61	GAAAGAGCTCTCATTTCTCCTCTTCTATTTTTTTTAAACCACTATGAGCGTTCTTAGAGAT 120	
121	CAAAAGGGGCTGTGGCATCTTATCAAGTTGGCCCAAGTCTGACCGTGATGGCGGCATT 180	
121	CAAAAGGGGCTGTGGCATCTTATCAAGTTGGCCCAAGTCTGACCGTGATGGCGGCATT 180	
181	GGCTTGGGCTTCCTACCTCGAGTTTCCGGAGACACACTGGAGAGCGTGGCCCTTCAAC 240	
181	GGCTTGGGCTTCCTACCTCGAGTTTCCGGAGACACACTGGAGAGCGTGGCCCTTCAAC 240	
241	CTCTTCATGCTGGCGCTTGTGTGACAGTGGGCAATCTGCTGACAGGCTTCTTAGCCAG 300	
241	CTCTTCATGCTGGCGCTTGTGTGACAGTGGGCAATCTGCTGACAGGCTTCTTAGCCAG 300	
301	TTCCCTTCTGGAGAGTGGTATCACTGTTCAAGTATTCGGCTGGCCACATGAGTCT 360	
301	TTCCCTTCTGGAGAGTGGTATCACTGTTCAAGTATTCGGCTGGCCACATGAGTCT 360	
361	TGTGTGGTCTATCTCACTGAGTCTGCTTTGGGGAAGTCAATTGGCCAGTTGGTG 420	
361	TGTGTGGTCTATCTCACTGAGTCTGCTTTGGGGAAGTCAATTGGCCAGTTGGTG 420	
421	GTGATGTGCTGTGAGAGTGAACGCTTTAGGCACTGAGAGTGTATCAGTAATATC 480	
421	GTGATGTGCTGTGAGAGTGAACGCTTTAGGCACTGAGAGTGTATCAGTAATATC 480	

OY	481	TTCAACACAGATATACCA	CAATGAACATGATGACATCTACGTTGTGGACGCCATTTTGGG	540
Db	481	TTCAACACAGATATACCA	CAATGAACATGATGACATCTACGTTGTGGACGCCATTTTGGG	540
OY	541	CTGTCTGTGGCTGTGGT	GCCTGGCCAAAGCCTCTAACCCGAGGAAACGAGATTAAGATCAG	600
Db	541	CTGTCTGTGGCTGTGGT	GCCTGGCCAAAGCCTCTAACCCGAGGAAACGAGATTAAGATCAG	600
OY	601	ACAGACAGATACCCAG	ATTGTCTGTGCATGCTGGGCGCCCTCTTTGTGTGATGTTCTGG	660
Db	601	ACAGACAGATACCCAG	ATTGTCTGTGCATGCTGGGCGCCCTCTTTGTGTGATGTTCTGG	660
OY	661	CCAAATTTCAACTGCT	CTGTGTGAGAGTCCAAATCGAAAGAAAGAACCCGTTTCAAC	720
Db	661	CCAAATTTCAACTGCT	CTGTGTGAGAGTCCAAATCGAAAGAAAGAACCCGTTTCAAC	720
OY	721	ACCTACTATGCTGTAG	CAAGTCAAGCTGTGAGACGATCTCAGGGTTCATCTTGTGCTCAC	780
Db	721	ACCTACTATGCTGTAG	CAAGTCAAGCTGTGAGACGATCTCAGGGTTCATCTTGTGCTCAC	780
OY	781	CCCCAAGGAGAGATCAG	CAAGACTTAATGTGCA CAGTCCGTTGTGGCAGAGAGCGTGGCT	840
Db	781	CCCCAAGGAGAGATCAG	CAAGACTTAATGTGCA CAGTCCGTTGTGGCAGAGAGCGTGGCT	840
OY	841	GTGGGTACTCTCGT	GTCACTGATCCCTTCTCCGTGGCTTGCATGATGCTGGGTCTTGTG	900
Db	841	GTGGGTACTCTCGT	GTCACTGATCCCTTCTCCGTGGCTTGCATGATGCTGGGTCTTGTG	900
OY	901	GCTGGGCTGATCTCG	TCCGTCGGGGGAGCCAAAGTACTGCCGGGGGTGTGTAAACCGAGTGGC	960
Db	901	GCTGGGCTGATCTCG	TCCGTCGGGGGAGCCAAAGTACTGCCGGGGGTGTGTAAACCGAGTGGC	960
OY	961	GGGATTTCCCA CAG	CTCCATCATGAGGCTTACAACCTTCAGCTTGCCTGGGTCTGCTTGGAGAG	1020
Db	961	GGGATTTCCCA CAG	CTCCATCATGAGGCTTACAACCTTCAGCTTGCCTGGGTCTGCTTGGAGAG	1020
OY	1021	ATCATCTACATTTGT	GCTGCTGTGCTTGTATACCGTGGAGCCGCGCAATGCAATGTGGC	1080
Db	1021	ATCATCTACATTTGT	GCTGCTGTGCTTGTATACCGTGGAGCCGCGCAATGCAATGTGGC	1080
OY	1081	TTCCAGAGCTCTCCT	CAGCACTTGGGGAACTGACCTTGGCCATGCTGATGACTCTCACGCT	1140
Db	1081	TTCCAGAGCTCTCCT	CAGCACTTGGGGAACTGACCTTGGCCATGCTGATGACTCTCACGCT	1140
OY	1141	GGTCTCTGTACAG	CTTGGTCTCTTAATCTTAAATATGAAAGCACTCATGAGCTAAA	1200
Db	1141	GGTCTCTGTACAG	CTTGGTCTCTTAATCTTAAATATGAAAGCACTCATGAGCTAAA	1200
OY	1201	TATTTTGTATGAC	CCAAAGTTTCTGTGAAAGTTTCTCATTTGGCTGTGGATTTTAA	1254
Db	1201	TATTTTGTATGAC	CCAAAGTTTCTGTGAAAGTTTCTCATTTGGCTGTGGATTTTAA	1254

RESULT 11

AF187846

LOCUS AF187846 1458 bp mRNA linear PRI 31-OCT-1999

DEFINITION Homo sapiens Rhd type IIa protein mRNA, complete cds.

ACCESSION AF187846

VERSION AF187846.1 GI:6164858

KEYWORDS FLI\_CDNA.

SOURCE Homo sapiens.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 1458)

AUTHORS Huang, C.H., Chen, Y. and Reid, M.

JOURNAL Human D (IIa) erythrocytes: Rhd protein is associated with multiple dispersed amino acid variations

MEDLINE Am. J. Hematol. 55 (3), 139-145 (1997)

PUBMED 97398395

REFERENCE 2 (bases 1 to 1458)

AUTHORS Huang, C.H.

Db 601 AGAGCAAGATACCCAGTTTGTCTCCATGCGGGCGCCCTCTTCTGTGATGTTCTGG 660  
Qy 661 CCAAGTTTCACTCTGCTCTGCTGAGAGTCCAAATCGAAAGAAAGAACCGCTGTTCAAC 720  
Db 661 CCAAGTTTCACTCTGCTCTGCTGAGAGTCCAAATCGAAAGAAAGAACCGCTGTTCAAC 720  
Qy 721 ACTACTATGCTGTAGAGTCAAGCTGTGTGACAGCCATCTCAGGGTCACTCTGAGCTAC 780  
Db 721 ACTACTATGCTGTAGAGTCAAGCTGTGTGACAGCCATCTCAGGGTCACTCTGAGCTAC 780  
Qy 781 CCCCAGGAGAAAGATCAGCAAGACTTATGTGACAGTGGGTGTGGCAGAGGCGTGGCT 840  
Db 781 CCCCAGGAGAAAGATCAGCAAGACTTATGTGACAGTGGGTGTGGCAGAGGCGTGGCT 840  
Qy 841 GTGGTACTGCTGTACCTATCCCTTCTCGTGGCTTGCATAGTCTGGTCTTGTG 900  
Db 841 GTGGTACTGCTGTACCTATCCCTTCTCGTGGCTTGCATAGTCTGGTCTTGTG 900  
Qy 901 GCTGGGCTGATCTCCGTGCGGGGAGCCAAAGTACTGCGGGGTGTGTAAACGAGTGTCTG 960  
Db 901 GCTGGGCTGATCTCCGTGCGGGGAGCCAAAGTACTGCGGGGTGTGTAAACGAGTGTCTG 960  
Qy 961 GGGATTTCCCAAGCTCATCATGTGGGCTAACCTTCAAGCTTGTGGGTCTGTTGAGAG 1020  
Db 961 GGGATTTCCCAAGCTCATCATGTGGGCTAACCTTCAAGCTTGTGGGTCTGTTGAGAG 1020  
Qy 1021 ATCATCTATGTCGTCTGCTGTGTATACCGTGGAGCCGGGAATGGCATGTATGGC 1080  
Db 1021 ATCATCTATGTCGTCTGCTGTGTATACCGTGGAGCCGGGAATGGCATGTATGGC 1080  
Qy 1081 TTCAGGCTCTCTCAGCATTTGGGGAACCTCAGCTTGGCATGTATGATCTCAAGCTCT 1140  
Db 1081 TTCAGGCTCTCTCAGCATTTGGGGAACCTCAGCTTGGCATGTATGATCTCAAGCTCT 1140  
Qy 1141 GGTCTCTGACAGGTTTGTCTCTAATCTTAAATATGGAAGAACCTCAGTAA 1200  
Db 1141 GGTCTCTGACAGGTTTGTCTCTAATCTTAAATATGGAAGAACCTCAGTAA 1200  
Qy 1201 TATTTGATGACCAAGTTTCTGGAAGTTTCTCATTTGGCTGTGATTTAA 1254  
Db 1201 TATTTGATGACCAAGTTTCTGGAAGTTTCTCATTTGGCTGTGATTTAA 1254

RESULT 12  
LOCUS S57971 1251 bp mRNA linear PRI 28-JUN-1993  
DEFINITION Rh polypeptide II (clone RHP1) [human, erythroid cells, mRNA  
S57971  
ACCESSION S57971  
VERSION S57971.1 GI:299051  
KEYWORDS  
SOURCE Homo sapiens erythroid cells.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 1251)  
AUTHORS Kaji, E., Umenishi, F., Iwamoto, S. and Ikemoto, S.  
TITLE Isolation of a new cDNA clone encoding an Rh polypeptide associated  
with the Rh blood group system  
JOURNAL Hum. Genet. 91 (2), 157-162 (1993)  
MEDLINE 93216282  
PUBMED 7916743  
REMARK GenBank staff at the National Library of Medicine created this  
entry [NCBI gi299051.1] from the original journal article.  
This sequence comes from Fig. 3.  
Map location: 1p34.3-36.1.  
FEATURES  
source  
1..1251  
Location/Qualifiers  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
1..1251  
/partial

CDS

4

/note="Rh polypeptide II; This sequence comes from Fig. 3;  
authors show full coding region except for the stop codon"  
/codon\_start=1  
/protein\_id="PAB26081.1"  
/db\_xref="GI:299051"  
/translation="MSSKTPRSVRRLPLMALTLLEAALLLPFFTHYDASLDQKEL  
VASVYQGDITVVAALIGLFLTSSFRHSSVAEVLPMALVQNALIDGLSOP  
PGKAVITLPSIRATMSALSVLISADAVKVLADLVVAVLEVALGLRLVLSNI  
FNTDYNNMHIIYVPAAYFGLTVAACLPRKLPETGDKQATIPSLISALGALFLM  
FMSFNSALLRSPIERKNAVENVYVAVSVTAISGSLAHPOGKISKTYVSAVLA  
GVAVTGSCHLIPSLPMLAMTGLVAGLISVGAKXYLPVCONRLGIHSSIMGYNLSL  
LGILGETIYVILVLDITVAGNENIGFOYLLISIGELSLAIVIALTSLGLTLNLKI  
WKAHPKAFYFDQDFWKPFLAVGP"  
BASE COUNT 240 a 322 c 351 g 338 t  
ORIGIN  
Query Match 98.9%; Score 1239.8; DB 9; Length 1251;  
Best Local Similarity 99.4%; Pred. No. 0;  
Matches 1244; Conservative 0; Mismatches 7; Indels 0; Gaps 0;  
Qy 1 ATGAGCTTAAGTACCGCGGTCTGTCCGGCGCTGCTGCCCCCTCTGGGCCCTTAACATG 60  
Db 1 ATGAGCTTAAGTACCGCGGTCTGTCCGGCGCTGCTGCCCCCTCTGGGCCCTTAACATG 60  
Qy 61 GAACAGCTCTCATTTCTCTCTTATTTTAAACCACTATGACGCTTCTTATGAGAT 120  
Db 61 GAACAGCTCTCATTTCTCTCTTATTTTAAACCACTATGACGCTTCTTATGAGAT 120  
Qy 121 CAAAGGGGCTGTGGCATCTTATCAAGTTGGCCAAATCGACCGGTATGGCGGCAT 180  
Db 121 CAAAGGGGCTGTGGCATCTTATCAAGTTGGCCAAATCGACCGGTATGGCGGCAT 180  
Qy 181 GGTGGGCTCTCTCACTCGAGTTTCCGAGACACAGCTGAGAGTGGCTTCAAC 240  
Db 181 GGTGGGCTCTCTCACTCGAGTTTCCGAGACACAGCTGAGAGTGGCTTCAAC 240  
Qy 241 CTCTTCATGCTGGCGCTTGTGTGTCAGTGGGCAATCTCTGTCAGCGCTTCTAAG 300  
Db 241 CTCTTCATGCTGGCGCTTGTGTGTCAGTGGGCAATCTCTGTCAGCGCTTCTAAG 300  
Qy 301 TTCCCTTGTGGAGGTGTATACACTGTTCAGTATGCGCTGGCCACCATAGTGT 360  
Db 301 TTCCCTTGTGGAGGTGTATACACTGTTCAGTATGCGCTGGCCACCATAGTGT 360  
Qy 361 TTGTGCGGCTGATCTCAAGTGTCTTGGGGAAGCTCACTTGGCCAGTGTGTG 420  
Db 361 TTGTGCGGCTGATCTCAAGTGTCTTGGGGAAGCTCACTTGGCCAGTGTGTG 420  
Qy 421 GTGATGTGTGTGTGAGGTGACAGCTTTAGGCAACCTGAGATGTATCATGATATC 480  
Db 421 GTGATGTGTGTGTGAGGTGACAGCTTTAGGCAACCTGAGATGTATCATGATATC 480  
Qy 481 TTCAACACAGCTACCAATGATGATGACATCTTACGTTTCCGAGCCCTATTTGGG 540  
Db 481 TTCAACACAGCTACCAATGATGATGACATCTTACGTTTCCGAGCCCTATTTGGG 540  
Qy 541 CTGTCTGTGGCTGTGTGCTCCAAAGGCTCTTACCCGAGGAAAGAGATTAAGATCAG 600  
Db 541 CTGTCTGTGGCTGTGTGCTCCAAAGGCTCTTACCCGAGGAAAGAGATTAAGATCAG 600  
Qy 601 ACAGCAAGATACCCAGTTTGTCTGCCATGTGTGGCGCCCTCTTGTGTGATGTTCTGG 660  
Db 601 ACAGCAAGATACCCAGTTTGTCTGCCATGTGTGGCGCCCTCTTGTGTGATGTTCTGG 660  
Qy 661 CCAAGTTTCACTCTGCTCTGCTGAGAGTCCAAATCGAAAGAAAGAACCGCTGTTCAAC 720  
Db 661 CCAAGTTTCACTCTGCTCTGCTGAGAGTCCAAATCGAAAGAAAGAACCGCTGTTCAAC 720  
Qy 721 ACTACTATGCTGTAGAGTCAAGCTGTGTGACAGCCATCTCAGGGTCACTCTGAGCTAC 780  
Db 721 ACTACTATGCTGTAGAGTCAAGCTGTGTGACAGCCATCTCAGGGTCACTCTTGGCTAC 780  
Qy 781 CCCCAGGAGAAAGATCAGCAAGACTTATGTGACAGTGGGTGTGGCAGAGGCGTGGCT 840

```

Db 781 CCCCCAGGAAGATGACAGACCTTATGTGCAAGTGGGCTGGTGGACGAGGCGTGCT 840
Qy 841 GTGGGATCCTGCTGTCACCTGATCCCTTCTCCGCGCTTGGCCATGAGTGGCTTGTG 900
Db 841 GTGGGATCCTGCTGTCACCTGATCCCTTCTCCGCGCTTGGCCATGAGTGGCTTGTG 900
Qy 901 GCTGGGCTGATCTCCGTCGGGGAGCCAAAGTACCTGCCGAGTGTGTAACCGAGTGTG 960
Db 901 GCTGGGCTGATCTCCGTCGGGGAGCCAAAGTACCTGCCGAGTGTGTAACCGAGTGTG 960
Qy 961 GGGATTCGCCACAGCTCATGATGGGCTACAACTTCCAGCTTGGCGGTCTGGTGGAGAG 1020
Db 961 GGGATTCGCCACAGCTCATGATGGGCTACAACTTCCAGCTTGGCGGTCTGGTGGAGAG 1020
Qy 1021 ATCATCTACATTTGCTGCTGCTGATACCGTCGGAGCGGCAATGAGCATTTGGC 1080
Db 1021 ATCATCTACATTTGCTGCTGCTGATACCGTCGGAGCGGCAATGAGCATTTGGC 1080
Qy 1081 TTCAGGTCCTCTCAGCATTTGGGAACTCAGCTTGGCCATGATGATGCTCTCAGCTCT 1140
Db 1081 TTCAGGTCCTCTCAGCATTTGGGAACTCAGCTTGGCCATGATGATGCTCTCAGCTCT 1140
Qy 1141 GGTCTCTGACAGGTTTGTCTTAAATCTTAAATATGGAAGACCTCATGAGGCTTAA 1200
Db 1141 GGTCTCTGACAGGTTTGTCTTAAATCTTAAATATGGAAGACCTCATGAGGCTTAA 1200
Qy 1201 TATTTGATGACCAAGTTTCTGGAAGTTTCTCATTTGGCTGTGATTT 1251
Db 1201 TATTTGATGACCAAGTTTCTGGAAGTTTCTCATTTGGCTGTGATTT 1251

```

```

RESULT 13
AB046420 1254 bp mRNA linear PRI 28-JUL-2001
LOCUS Homo sapiens RHD1 mRNA for Rh blood group D antigen, complete cds.
DEFINITION AB046420
ACCESSION AB046420
VERSION AB046420.1 GI:12381901
KEYWORDS
SOURCE Homo sapiens
ORGANISM Homo sapiens cDNA to mRNA.
Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

```

```

REFERENCE
AUTHORS 1 Toshinori, O., Junko, T., Yoshihiko, T. and Eiji, K.
TITLE Homo sapiens RHD1 mRNA for Rh blood group antigen Rhd, complete
JOURNAL cds
REFERENCE 2 Published only in Database (2001)
AUTHORS 2 (bases 1 to 1254)
TITLE Toshinori, O., Junko, T., Yoshihiko, T. and Eiji, K.
JOURNAL Direct Submission
Submitted (21-JUL-2000) Omi Toshinori, Jichi Medical School, Dept.
of Legal Medicine and Human Genetics, 3311-1 Yakushiji
Mitsukawauchi, Kawachi, Tochigi 329-0498, Japan
E-mail: t-omi@jichi.ac.jp, URL: www.jichi.ac.jp, Tel: 81-285-58-7342,
Fax: 81-285-44-4902

```

```

FEATURES
source
1..1254
Location/Qualifiers

```

```

gene
1..1254
/organism="Homo sapiens"
/db_xref="taxon:9606"
1..1254
/gene="RHD1"
1..1254
/gene="RHD1"
/codon_start=1
/product="Rh blood group D antigen"
/protein_id="BAB21261.1"
/db_xref="GI:12381902"
/translacion="MSKYPKSRVRCPLWALTEALILLFFETHYDLEOKSL
VASVOGLTVAALIGLFTSFRHSVSAFNLPMALGVOMAILDGLPSOP
SKRVLTLSRLATMSALSVLSVDVLKRNLAQVNVVLEVTALGMLRVYISNI
PNDYHNNMHHIVFAIVFGLSVAMCLPKPLPSTEDKDDTATIPSLISANLGLFLWM
FWSVNSPLIRSPIORNMENNTYALAVSVTAISSSLAHPRKISMTYVSAVLA

```

```

GGVAVGTSCHLIPSPWLANVLGVAGLISVGAKYIPGCCNRVLGIPHSINGNPSL
LAGLIGLIVLVLDVAGNMGIFQVYLSTIGELSLAIVIALTGLTLNLT
WRAHEAKTFDDVFWKFPFLAVGF"
BASE COUNT 242 a 323 c 347 g 342 t
ORIGIN
Query Match 98.9%; Score 1239.6; DB 9; Length 1254;
Best Local Similarity 99.3%; Pred. No. 0;
Matches 1245; Conservative 0; Mismatches 9; Indels 0; Gaps 0;
Qy 1 ATGAGCTCTAAGTACCGCGGCTGTGTCGGGCTGCGCTGCGCCCTGAGCCCTTAACCTG 60
Db 1 ATGAGCTCTAAGTACCGCGGCTGTGTCGGGCTGCGCTGCGCCCTGAGCCCTTAACCTG 60
Qy 61 GAAGCACTCTCATCTCTCTTCTATTTTAAACCACTATGACGCTCTTGAAGAT 120
Db 61 GAAGCACTCTCATCTCTCTTCTATTTTAAACCACTATGACGCTCTTGAAGAT 120
Qy 121 CAAAAGGGGCTCGTGCATCCATCAAGTTGGCCAGATCTGACCTGATGGCGCAT 180
Db 121 CAAAAGGGGCTCGTGCATCCATCAAGTTGGCCAGATCTGACCTGATGGCGCAT 180
Qy 181 GCGTTGGGCTTCCCTACCTGAGTTTCCGAGACACAGCTGAGACATGTGGCCTTGAAC 240
Db 181 GCGTTGGGCTTCCCTACCTGAGTTTCCGAGACACAGCTGAGACATGTGGCCTTGAAC 240
Qy 241 CTCTTCATGCTGCGCTTGGTGTGTCAGTGGGCAATCTGCTGACGCGCTTCTAGACAG 300
Db 241 CTCTTCATGCTGCGCTTGGTGTGTCAGTGGGCAATCTGCTGACGCGCTTCTAGACAG 300
Qy 301 TTCCTTCTGGGAAGTGTGTCATCACTGTCAGTATTCGGCTGGCCACCATGATGCT 360
Db 301 TTCCTTCTGGGAAGTGTGTCATCACTGTCAGTATTCGGCTGGCCACCATGATGCT 360
Qy 361 TTGTCGCTGTCATGTCATGAGTGCATGCTGCTTGGGGAAGTCACTTGGCGCACTGTG 420
Db 361 TTGTCGCTGTCATGTCATGAGTGCATGCTGCTTGGGGAAGTCACTTGGCGCACTGTG 420
Qy 421 GTGATGTGCTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 480
Db 421 GTGATGTGCTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 480
Qy 481 TTCAACACAGACTACCAATGATGATGATGATGATGATGATGATGATGATGATGATG 540
Db 481 TTCAACACAGACTACCAATGATGATGATGATGATGATGATGATGATGATGATGATG 540
Qy 541 CTGCTGTGCTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 600
Db 541 CTGCTGTGCTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 600
Qy 601 ACAGCAAGATACCCAGTTTGTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 660
Db 601 ACAGCAAGATACCCAGTTTGTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 660
Qy 661 CCAAGTTTCAACTCTGCTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 720
Db 661 CCAAGTTTCAACTCTGCTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 720
Qy 721 ACTTACTATGTCGTGACATGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 780
Db 721 ACTTACTATGTCGTGACATGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 780
Qy 781 CCCCAGGGAAGATCAGACAGACTTATGTCACAGTCGCGGTTGGAGAGAGGCTGCT 840
Db 781 CCCCAGGGAAGATCAGACAGACTTATGTCACAGTCGCGGTTGGAGAGAGGCTGCT 840
Qy 841 GTGGGATCCTGCTGTCACCTGATCCCTTCTCCGCTTGGCCATGAGTGGCTTGTG 900
Db 841 GTGGGATCCTGCTGTCACCTGATCCCTTCTCCGCTTGGCCATGAGTGGCTTGTG 900
Qy 901 GCTGGGCTGATCTCCGTCGGGGAGCCAAAGTACCTGCCGAGTGTGTAACCGAGTGTG 960
Db 901 GCTGGGCTGATCTCCGTCGGGGAGCCAAAGTACCTGCCGAGTGTGTAACCGAGTGTG 960

```







Db 1081 TTCAGGCTCTCTCCACATCTGGGAACTCAGCTGGCCATCGGATAGCTCTACGCTCT 1140  
QY 1141 GGTCCTCGACAGGTTGGCTCTTAATCTTAAATATGAAAGACCTCATGAGGCTAAA 1200  
Db 1141 GGTCCTCGACAGGTTGGCTCTTAAATCTTAAATATGAAAGACCTCATGAGGCTAAA 1200  
QY 1201 TATTTGATGACCAAGTTTCTGGAAGTTTCTCATTTGGCTGTGGATTTTAA 1254  
Db 1201 TATTTGATGACCAAGTTTCTGGAAGTTTCTCATTTGGCTGTGGATTTTAA 1254

RESULT 15  
AF037626 1247 bp mRNA linear PRI 05-JAN-1999  
LOCUS Homo sapiens truncated RHD (RHD mutant) mRNA, complete cds.  
DEFINITION AF037626  
ACCESSION AF037626.1 GI:4104640  
VERSION AF037626.1 GI:4104640  
KEYWORDS  
SOURCE Homo sapiens.  
ORGANISM Homo sapiens.  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
REFERENCE 1 (bases 1 to 1247)  
AUTHORS Andrews,K.A., Walter,L.C., Saul,A. and Hyland,C.A.  
TITLE The Rn D antigen negative trait in an Rn Ccpe phenotypic Caucasian  
attributed to a four nucleotide deletion in the RH D gene  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 1247)  
AUTHORS Andrews,K.A. and Hyland,C.A.  
TITLE Direct Submission  
JOURNAL Submitted (10-DEC-1997) Malaria and Arbovirus Unit, Queensland  
Institute of Medical Research, 300 Herston road, Herston, QLD 4006,  
Australia

FEATURES  
source Location/Qualifiers  
1..1247  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="1"  
/map="1p34.3-p36.1"  
1..1247  
/gene="RHD"  
/note="RHD gene with a four nucleotide deletion at the  
start of exon 4, resulting in a D antigen negative  
phenotype"  
1..498  
/gene="RHD"  
/codon\_start=1  
/product="truncated Rhd"  
/protein\_id="A4D02100.1"  
/db\_xref="GI:4104641"  
/translation="MSKVPKSVRRCPLMALTLLEALLILFFETHYDASLDOKL  
VASIQVQDITVMAATIGLGLTSSFRHSSVAENFMALGLVOMAILDGLTSLQEP  
SGKAVITLFSIRLMTSALSVLISVDVILKVMLOLVNVLVEVTAIGMLRMVLSNI  
FNTT"  
BASE COUNT 237 a 322 c 349 g 339 t  
ORIGIN

Query Match 98.2%; Score 1231.4; DB 9; Length 1247;  
Best Local Similarity 99.6%; Pred. No. 0;  
Matches 1246; Conservative 0; Mismatches 1; Indels 4; Gaps 1;

QY 1 ATGAGCTTAAGTACCGCGGTGTCTCGGCGCTGCGCCCTCTGGGCGCTTAACACTG 60  
Db 1 ATGAGCTTAAGTACCGCGGTGTCTCGGCGCTGCGCCCTCTGGGCGCTTAACACTG 60  
QY 61 GAAGCAGCTCTCATTTCTCTCTTATTTTAAACCACTATGACGCTTCTTAGAGAT 120  
Db 61 GAAGCAGCTCTCATTTCTCTCTTATTTTAAACCACTATGACGCTTCTTAGAGAT 120  
QY 121 CAAAGGGGCTGTGGCATCTCTATCAAGTTGGCCAAAGTCTGACCGGTATGGCGGCACTT 180  
Db 121 CAAAGGGGCTGTGGCATCTCTATCAAGTTGGCCAAAGTCTGACCGGTATGGCGGCACTT 180

QY 181 GGGTGGGCTTCTCTCACTCGAGTTTCCGAGACACAGCTGAGCAAGTGTGGCTTCAAC 240  
Db 181 GGGTGGGCTTCTCTCACTCGAGTTTCCGAGACACAGCTGAGCAAGTGTGGCTTCAAC 240  
QY 241 CTCTTCAATGCTGGCGCTTGGTGTGCAAGTGGCAATCTCTGTGAAGGCTTCTGAGCCAG 300  
Db 241 CTCTTCAATGCTGGCGCTTGGTGTGCAAGTGGCAATCTCTGTGAAGGCTTCTGAGCCAG 300  
QY 301 TTCCCTTCTGGGAAAGGTGATCACTGTGAGTATTTGGGCTGGGCAACATGAGTGT 360  
Db 301 TTCCCTTCTGGGAAAGGTGATCACTGTGAGTATTTGGGCTGGGCAACATGAGTGT 360  
QY 361 TTGTGCGTGTGATCTTCAGTGAATGCTGTCTTGGGGAAGGCTCAACTTGGCGAGTTGGTG 420  
Db 361 TTGTGCGTGTGATCTTCAGTGAATGCTGTCTTGGGGAAGGCTCAACTTGGCGAGTTGGTG 420  
QY 421 GTGATGGTCTGTGGAGTGAAGTGAACAGTTTAGCAACTGAGATGTCTATCATATATC 480  
Db 421 GTGATGGTCTGTGGAGTGAAGTGAACAGTTTAGCAACTGAGATGTCTATCATATATC 480  
QY 481 TTCAACACAGACTACCAATGAACATGATGACATCTACGTTTCCGAGCTATTTTGGG 540  
Db 481 TTCAACACAGACTACCAATGAACATGATGACATCTACGTTTCCGAGCTATTTTGGG 540  
QY 541 CTGTCTGTGGCTGTGGTCTGCGCAAGGCTCTACCCGAGGGAACGAGATGAAGATCAG 600  
Db 541 CTGTCTGTGGCTGTGGTCTGCGCAAGGCTCTACCCGAGGGAACGAGATGAAGATCAG 600  
QY 601 ACAGCAACGATACCCAGTTTGTCTGCGCATCTGGGCGGCTCTTCTTGTGATGTCTGG 660  
Db 601 ACAGCAACGATACCCAGTTTGTCTGCGCATCTGGGCGGCTCTTCTTGTGATGTCTGG 660  
QY 661 CCAAGTTCAACTCTGCTGCTGAGAGTCCAAATGCAAGGAAGATGCGGTTCAC 720  
Db 661 CCAAGTTCAACTCTGCTGCTGAGAGTCCAAATGCAAGGAAGATGCGGTTCAC 720  
QY 721 ACCTATATCTGTGAGCAGTCAAGGCTGTGAGACAGCCATCTCAGGGTCACTTGGCTCAC 780  
Db 721 ACCTATATCTGTGAGCAGTCAAGGCTGTGAGACAGCCATCTCAGGGTCACTTGGCTCAC 780  
QY 781 CCCCAGGGAAGATCAGCAAGCTTATATGTCACAGTGGGTGTGGAGAGCGTGGCT 840  
Db 781 CCCCAGGGAAGATCAGCAAGCTTATATGTCACAGTGGGTGTGGAGAGCGTGGCT 840  
QY 841 GTGGGTACCTGTGTACCTGATTCCTTCCGCTGGCTTGCATGATGTCTGGCTGTG 900  
Db 841 GTGGGTACCTGTGTACCTGATTCCTTCCGCTGGCTTGCATGATGTCTGGCTGTG 900  
QY 901 GCTGGGCTGATCTCGTGGGAGGAGCAAGTACCTGCGGGGTGTGTAACCGAGTCTG 960  
Db 901 GCTGGGCTGATCTCGTGGGAGGAGCAAGTACCTGCGGGGTGTGTAACCGAGTCTG 960  
QY 961 GGGATTTCCCAACAGCTTCATATGAGGCTACCACTTCACTGCTGGCTGTGGAGAG 1020  
Db 961 GGGATTTCCCAACAGCTTCATATGAGGCTACCACTTCACTGCTGGCTGTGGAGAG 1020  
QY 1021 ATCATCTAATGTGCTGCTGCTGATACCGTGGAGCCGGCAATGATGATGGC 1080  
Db 1021 ATCATCTAATGTGCTGCTGCTGATACCGTGGAGCCGGCAATGATGATGGC 1080  
QY 1081 TTCCAGGCTCTCTCAGCATTTGGGGAAGTCAAGCTTGGCCATCGGATAGCTTCAAGCTCT 1140  
Db 1081 TTCCAGGCTCTCTCAGCATTTGGGGAAGTCAAGCTTGGCCATCGGATAGCTTCAAGCTCT 1140  
QY 1141 GGTCCTCGACAGGTTGGCTCTTAATCTTAAATATGAAAGACCTCATGAGGCTAAA 1200  
Db 1141 GGTCCTCGACAGGTTGGCTCTTAATCTTAAATATGAAAGACCTCATGAGGCTAAA 1200  
QY 1201 TATTTGATGACCAAGTTTCTGGAAGTTTCTCATTTGGCTGTGGATT 1251  
Db 1201 TATTTGATGACCAAGTTTCTGGAAGTTTCTCATTTGGCTGTGGATT 1251

Tue Apr 8 11:24:31 2003

us-09-600-714-41.rge

Page 16

Search completed: April 8, 2003, 05:00:48  
Job time : 3630 secs

---